20

Location modelling and GIS

R L CHURCH

Location modelling involves the search for the best location of one or more facilities to support some desired function. Examples range from retail site location to the location of multiple ambulance dispatch points. GIS has played a large role in the siting of single facilities, including rights-of-way for roads and transmission lines. Its use in multi-facility location problems is a relatively recent development. This chapter describes some of the history of location search as supported by GIS. It also discusses some of the current impediments to the application of location models, issues associated with the integration of location models into GIS, and future needs in GIS functionality to support location models.

1 WHAT IS A LOCATION PROBLEM?

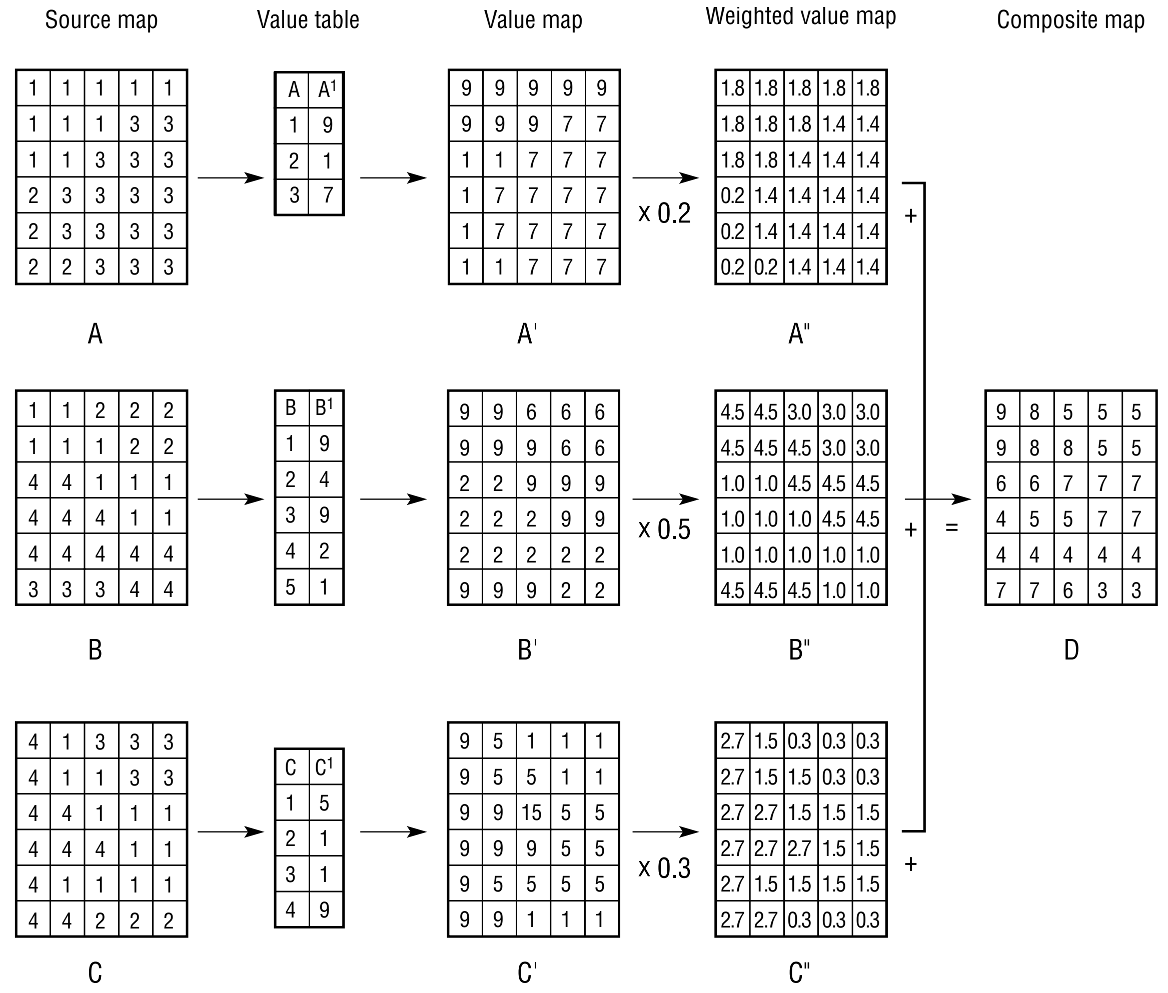
Webster's Dictionary defines location as (1) position in space, and (2) an area marked off for a specific purpose. Hence, a location problem would involve identifying a specific position or place for a specific function or activity. There are several common types of location problem associated with GIS. The most common involves the measurement of where something exists. Given adequate time, it is possible to measure the location of virtually anything on the Earth. This type of problem can be called a location measurement problem. The second type involves the search for an appropriate location for an activity. This problem can be called the locational search problem. It is common to refer to the locational search problem as a facility location problem. This type of problem can involve the placement of one activity (e.g. a retail store) or the placement of a set of interrelated facilities (e.g. fire stations to serve an urban area). Such problems are called single and multi-facility location problems. Other chapters in this book provide examples in the contexts of business and service planning (Birkin et al, Chapter 51), transportation (Waters, Chapter 59), and electoral districting (Horn, Chapter 67).

2 EARLY HISTORY AND DEVELOPMENT

Both locational measurement and locational search problems are important. The focus of this chapter will be on the problem of locational search. One of the early applications which helped to start the development of GIS involved locational search problems over large regions. In the 1970s many states began the development of geographical databases (in raster format) for the storage and retrieval of environmental and planning data. Examples of such systems include LUNR, a land-use and natural resource inventory of the State of New York, and the MAGI (Maryland Automated Geographic Information) database of Maryland. As an historical side note, Dangermond of ESRI (Environmental Systems Research Institute) was a principal developer of the MAGI database. During the mid 1970s there was considerable debate as to whether the electrical utilities could keep up with the increasing demand for electricity. Further, many questioned whether enough suitable sites for power plants existed. Out of this concern, the State of Maryland funded the Maryland Power Plant Siting Project and the development of a simple grid-based GIS. The State was divided into approximately

30 000 cells, each measuring 91.8 acres. In each cell 52 variables were measured, including land-use, land cover, proximity to a water source, seismicity etc. Each of these variables was converted to a subscore function and added together to form a site suitability score (Dobson 1979). From this, a map was produced which identified those cells which scored higher than one standard deviation above the mean. This process is depicted in Figure 1. The Planning Office of the State of Maryland still manages and updates this database, although it is now stored in both raster and vector format.

Another location problem which has evolved with GIS is the location of rights-of-way for roads and transmission power lines. It is interesting to note that much of the development of computerised modelling for such linear facilities has been coordinated with some type of GIS, principally raster. McHarg (1969) in *Design with Nature* developed a process which represents a precursor to the classic overlay process in GIS. McHarg's process involved developing a colour acetate map sheet for each basic theme that is used in the location of a corridor. The shading of a given map was from light to dark (high suitability to low suitability). When all of the acetate sheets were laid on top of each other, and placed over a light-table, the areas that had little colour were considered to be the best in terms of suitability. McHarg's process called for the tracing of the most direct route connecting the desired terminus points and which travelled through as much lightly coloured area as possible. Most



A

1	1	1	1	1
1	1	1	3	3
1	1	3	3	3
2	3	3	3	3
2	3	3	3	3
2	2	3	3	3

A	A^1
1	9
2	1
3	7

A'

9	9	9	9	9
9	9	9	7	7
1	1	7	7	7
1	7	7	7	7
1	7	7	7	7
1	1	7	7	7

A"

1.8	1.8	1.8	1.8	1.8
1.8	1.8	1.8	1.4	1.4
1.8	1.8	1.4	1.4	1.4
0.2	1.4	1.4	1.4	1.4
0.2	1.4	1.4	1.4	1.4
0.2	0.2	1.4	1.4	1.4

B

1	1	2	2	2
1	1	1	2	2
4	4	1	1	1
4	4	4	1	1
4	4	4	4	4
3	3	3	4	4

B	B^1
1	9
2	4
3	9
4	2
5	1

B'

9	9	6	6	6
9	9	9	6	6
2	2	9	9	9
2	2	2	9	9
2	2	2	2	2
9	9	9	2	2

B"

4.5	4.5	3.0	3.0	3.0
4.5	4.5	4.5	3.0	3.0
1.0	1.0	4.5	4.5	4.5
1.0	1.0	1.0	4.5	4.5
1.0	1.0	1.0	1.0	1.0
4.5	4.5	4.5	1.0	1.0

D

9	8	5	5	5
9	8	8	5	5
6	6	7	7	7
4	5	5	7	7
4	4	4	4	4
7	7	6	3	3

C

4	1	3	3	3
4	1	1	3	3
4	4	1	1	1
4	4	4	1	1
4	1	1	1	1
4	4	2	2	2

C	C^1
1	5
2	1
3	1
4	9

C'

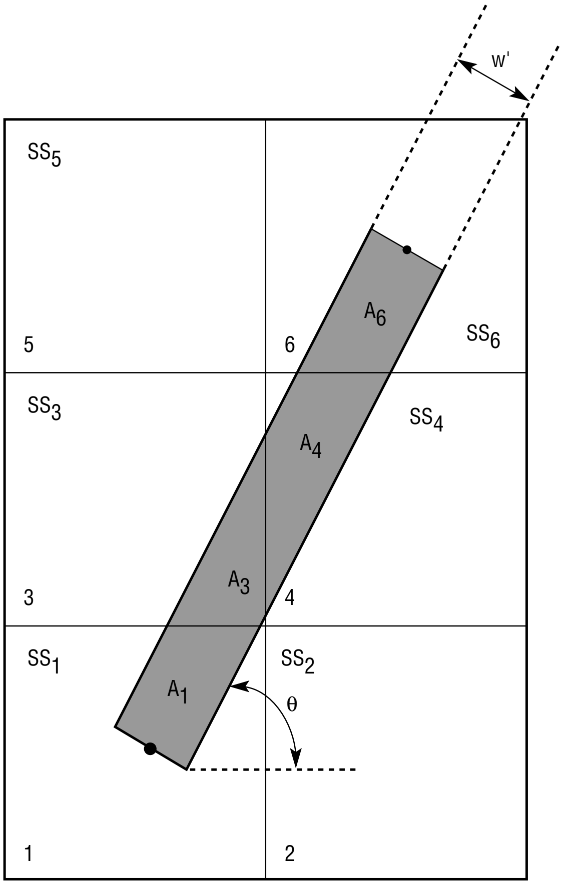
9	5	1	1	1
9	5	5	1	1
9	9	15	5	5
9	9	9	5	5
9	5	5	5	5
9	9	1	1	1

C"

2.7	1.5	0.3	0.3	0.3
2.7	1.5	1.5	0.3	0.3
2.7	2.7	1.5	1.5	1.5
2.7	2.7	2.7	1.5	1.5
2.7	1.5	1.5	1.5	1.5
2.7	2.7	0.3	0.3	0.3

Fig 1. Combining data layers and attributes in order to determine site suitability scores.

subsequent corridor routing applications have involved raster-based systems. Information associated with various attributes is typically converted to subscores and combined into suitability scores for compatibility with use for a corridor. Next, a network is defined where the nodes represent the centroids of the cells. This is depicted in Figure 2. The arcs depict possible directions which can be taken by the corridor through a given cell. Goodchild (1977) and Huber and Church (1985) have shown that considerable error may exist in the unneeded elongation of a route based upon the limits made in the compass directions an arc may take from a given centroid. Arc costs or weights are usually defined as some weighted combination of the suitability scores through which the arc passes. In Figure 3, a case is depicted where the arc has a width of influence, and the weights are associated with the footprint of the corridor in a given cell. Given a network, complete with arcs and associated traverse costs, a shortest path algorithm can be applied in order to identify the least cost (most suitable) route across the landscape. The beauty of this process is that it can be solved optimally for large problems. An example is given in Figure 4, which involves a landscape of more than 16 000 cells and 256 000 arcs and where the solution time is but a few minutes on even modest workstations. Because such a process is fast, it can be repeated for a wide variety of parameter weights associated with different components of the suitability score, allowing the user to test the sensitivity of the routing to various levels of importance weights. The more efficient a solution algorithm can be made for a given application, the more flexibility there is in designing a user interface to allow for sensitivity analysis and general spatial exploration (Church et al 1992).



Arc value = $A_1{}^*SS_1+A_3{}^*SS_3+A_4{}^*SS_4+A_6{}^*SS_6$

Fig 3. Determining the weighted cost for an arc.

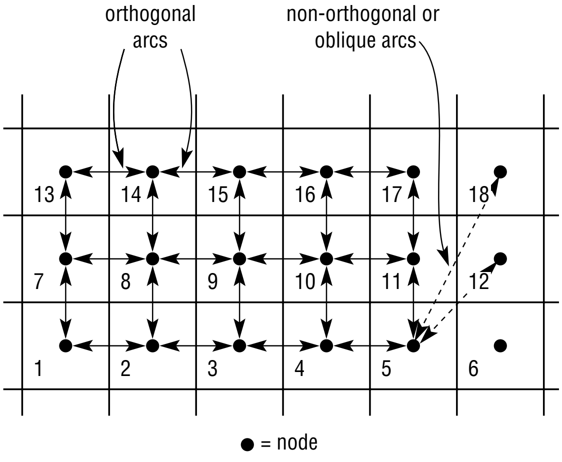


Fig 2. Depicting arc directions on a raster system.

Single-site location analysis can often be approached with some form of ranking or scoring process, which is based upon a set of attributes. From a landscape of potential site scores, areas are screened or filtered out if their scores are too low to be acceptable (as in the power plant siting example). This process is generally well supported in a number of GIS products. Documented applications abound in GIS conference proceedings associated with the use of such a site search process. In fact a number of companies provide data and services to assist in such general site searches. Further, special scoring processes like the Analytic Hierarchy Process have been linked to GIS in systems like Idrisi (1996).

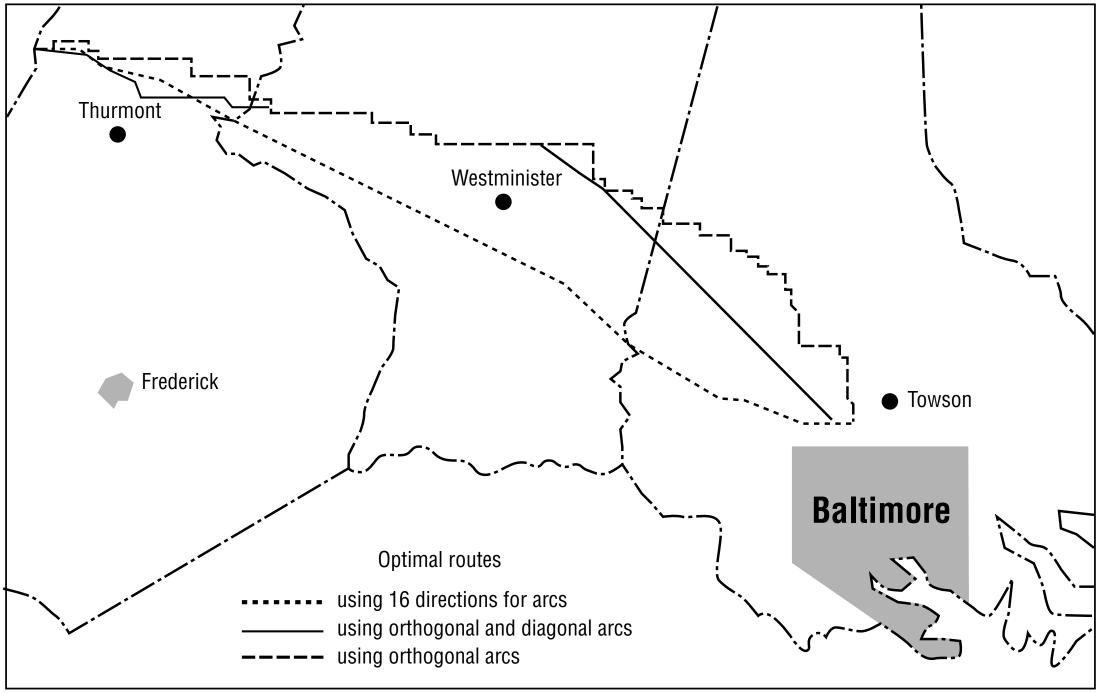


Fig 4. An efficient route using State of Maryland data.

Because such site searches can be performed with a modest set of GIS functions, it is both well supported in software and widely applied.

Beyond the search for an appropriate site or the search for a corridor across a landscape, a third major location search problem exists. This third major category represents simultaneous multiple-site selection. To explain this, consider the placement of ambulances for the purpose of emergency response. One common approach to developing a deployment plan is called System Status Management (SSM; Stout 1989). A deployment plan for SSM is established for each hour of the week. For example, the City of Denver has an SSM plan for the time interval of Tuesday 2–3 am. This plan is based on providing a specific number of ambulances in order to meet the expected demand with a very high probability (for example to provide 95 per cent response within 8 minutes). It also details how ambulances should be positioned during that hour, when all are available, when all but one are available, when all but two are available, etc. As ambulances are called for service during that hour, ambulances are repositioned to provide the best coverage possible according to the SSM plan. When ambulances are finished with a call, some or all are repositioned in order to deploy better what is currently available. To develop an SSM for this one hour alone, it is necessary to identify a location plan for *p* ambulances, *p*–1 ambulances, *p*–2 ambulances, etc. To do this requires the solution of a series of location problems, each involving the location of a multiple number of units (facilities). The multi-facility location problem will be discussed at greater length in the next several sections of this chapter.

3 DEVELOPING A FRAMEWORK FOR LOCATION MODEL CLASSIFICATION AND GIS

Most location models are variants of four general classes: median, covering, capacitated, and competitive. A median model involves locating a fixed number of facilities in such a manner that the average distance from any user to their closest facility is minimised. Covering models involve locating facilities in order to cover all or most demand within some desired service distance (often called the maximum service distance). The idea is that the more users who are served relatively close to a facility, the better the service. For example, in ambulance deployment a common goal is to serve at least 90 per cent of the population within 8 minutes. Classic median models are based upon the assumption that there are enough resources at each facility to handle whatever demand needs to be served. Thus, everyone is assumed to be served by their closest facility. Capacitated facility models place some limit on what can be accomplished at each facility (e.g. the number of units that can be manufactured, the amount of demand that can be served or assigned, the volume of garbage that can be handled per day etc.). Finally, competition models involve the case where a competitor has the capability to readjust to any location decisions other competitors make over predefined time frames. If you locate a new branch which exploits a poorly served area of your competitor, your competitor may over a period of time relocate a branch or locate a new one to recoup some of that lost market. Thus, in making a location decision, it becomes necessary to analyse what response your competitor may make to keep from a loss of business. Median, covering, and capacitated models are usually addressed as classical optimisation models, whereas competition models are often addressed by game theory and simulation. Good reviews of location research can be found in the literature (Beaumont 1981; Brandeau and Chiu 1989; Eiselt 1992; ReVelle 1987; Schilling et al 1993).

How a location problem is defined involves the spatial relationships between what is defined as a demand and what is defined as a facility. Even though demand is often spread across space (e.g. a neighbourhood, a census tract, or an apartment building), it is often represented as a single point. In fact, most models are based upon the assumption that demands are represented as points, although there are some notable exceptions to this general rule in the location science literature (see Wesolowsky and Love 1971, where demand is represented by continuous rectangular areas). Facility sites can be defined as points, lines, or areas. For example, in a corridor location problem, the facility usually represents a curvilinear facility like a roadway that connects two prespecified terminus points.

For continuous surface problems, facility locations are often allowed to be anywhere. In terms of network models, facility sites are most often described as nodes although considerable theory has been developed concerning the location along arcs or links as well. More often than not, both demand and facility sites are represented as discrete points since most solution algorithms have been developed for such cases. Miller (1996) has presented a classification of location models based upon the geometric representation of demand (or clients) and the geometric representation of facilities. For example, demands represented as polygons and facilities represented by polygons can be used to define a polygon–polygon location problem involving the location of a set of polygons to serve a set of weighted polygons. An example of such a problem can be found in very large-scale integration chip design and in production layout. Miller argues that GIS allows for better opportunities to represent location model features, such as facility size and shape, and that such improvements could have the potential to increase the relevance and flexibility of facility location models. Should this prove to be true, then GIS will become an integral part of many location model approaches in the future.

4 INTRODUCTION TO BASIC PROBLEMS IN THE IMPLEMENTATION OF LOCATION MODELS WITH GIS

Most single-site location search problems are solved by enumerating all possibilities. Even though there may be many possible sites, scoring and selecting the top sites is not a very computationally burdensome task (see Dobson 1979 for a classic application of site search in GIS). In contrast, the task of simultaneously selecting a configuration of sites to accomplish some objective is, indeed, significant. Recall from the above discussion that the City of Denver might have up to ten ambulances in operation at a given hour. Suppose 200 good deployment locations for ambulances have been identified in the city. This means that there are:

$$\binom{200}{10} = \frac{200!}{(10!)\,(200 - 10!)} = 2.245 \times 10^{16}$$

possible configurational plans (i.e. the number of distinct ways 200 sites can be selected ten at a time) that involve siting exactly ten ambulances. This is obviously a large number. In fact, this number is so large it is impossible to enumerate and evaluate each configuration within the lifetime of a given computer. Thus, enumerating all possible configurations and selecting the best solution is not a feasible approach. Consequently, the problem needs to be solved through the development of some type of algorithm or heuristic. To approach this type of problem, it is common first to develop a mathematical formulation for the problem. Then, solution techniques are designed to solve the problem and take advantage of its structure. This is described in the next section at greater length for two very popular location model constructs.

Many of the multi-facility location models are considered to be computationally complex (Garey and Johnson 1979). Beyond the computational task involved in solving a location model, there are several important issues related specifically to model integration and GIS:

1. *Compatibility of data structures.* There can be differences between the data structure that has been designed to best support a location model algorithm and the principal data model used in a GIS (Densham and Rushton 1991). As an example, to support a location heuristic in ARC/INFO, designers have an intermediate process create a data string that is read and used by the solution process. That is, the GIS data structure is not used directly by the solution process. Hence, the solution process and the GIS are loosely coupled by the exporting and importing of data and results.
2. *Demand representation and site identification.* In order to search for a solution to most location models, there must be a defined set of demand areas (represented by points, lines, or areas) as well as a set of predefined sites. GIS can serve an important role in the definition of demand areas and facility sites. A number of constructs are possible (as described in the previous section), although demand areas and facility sites are often represented as points. GIS can add significant value to a given application when data gathered for another purpose are available to characterise demand or the feasibility of specific sites or regions.
3. *Aggregation.* Before GIS datasets were available for location study, problem size was relatively small. Since information is characterised in a rather exact way in GIS, this detail has begun to be represented in site and demand definition. For example, in the past it might be common to represent a city by census tracts alone. Now, even enumeration districts or possibly block faces can be used to characterise and represent demand. In one recent application by a public utility, individual houses represent the basic unit of demand. Thus, the demand surface is often characterised to a relatively fine level of detail by hundreds if not thousands of demand points. The same can be true for the representation of potential facility sites (whether they be areas, points, or line segments). Since many solution approaches cannot easily handle thousands of demand points and sites, some type of data aggregation is necessary. Goodchild (1979) was the first to demonstrate that data aggregation can have a great effect in the absolute location of specific facilities. Issues and properties of aggregation have been addressed by numerous researchers in geographical analysis (e.g. Current and Schilling 1987, 1990; Hillsman and Rhoda 1978). The current study of aggregation techniques can be aided by use of GIS.
4. *Error propagation.* Even though error in data is a fact of life, it is important to understand how data error may propagate in any procedure that is used, e.g. aggregation. Further, a specific aggregation scheme may introduce error in problem representation. Veregin (Chapter 12; 1995) has analysed the propagation of error in GIS analysis for classical GIS functions like overlay (and see Heuvelink, Chapter 14; Openshaw and Alvanides, Chapter 18). How error propagates and the extent that it may alter results in location models is in need of further research. It is necessary to analyse such occurrences, as there can be diminished reliability in the final results.
5. *Visualisation.* Exploring and comparing results from a location model can be enhanced by visual aids (Densham 1994). This much is known. Just what types of visualisation scheme are the best is still subject to research (Arentze et al 1996). One popular visualisation is the presentation of a map with demand allocated to located facilities by directed lines. This ‘spider-like’ plot provides an easy-to-understand view of where the facilities are, which demand is served by which facility (i.e. service regions), as well as the potential area differences in service regions (Armstrong et al 1992). There is a major need to identify which possible views aid an analyst in generating and searching for the best solution. This is also an element in which GIS can be of significant value in location modelling, as many attributes can be presented simultaneously with model results. For example, assume a set of cellular telephone communication sites have been located. With elevation data, the location pattern can be draped on a surface depicting elevation (see Fry, Chapter 58).

5 INTEGRATING TWO CLASSIC LOCATION MODELS INTO GIS: P-MEDIAN AND MAXIMAL COVERING

The *p*-median location model involves the location of a fixed number *p* of facilities. The objective is to locate the *p* facilities in such a manner that the total weighted distance of serving all demand is minimised. Weighted distance for a demand point represents the amount of demand multiplied by the distance to the closest facility. For example, if demand is measured in terms of the number of trips that need to be made by users of the facility, then weighted distance represents the total mileage involved in going to the facility. For a fixed level of demand, minimising total weighted distance is equivalent to minimising average distance. This model form can address many different types of application, from locating schools and health clinics to locating road maintenance garages and emergency response vehicles. Because this model captures the essence of locating a set of facilities to serve an area by maximising accessibility, it has become a popular model for application. This model can be formulated mathematically using the following notation (ReVelle and Swain 1970):

i = index of demand areas or nodes,
p = the number of facilities to be located,
j = index of potential facility sites,
a_i = the amount of demand at area/node i
d_{ij} = the shortest distance between demand i and facility site j,

$$x_{ij} = \begin{cases} 1, \text{ if demand at } i \text{ assigns to a facility at } j \\ 0, \text{ otherwise} \end{cases}$$

$$x_{jj} = \begin{cases} 1, \text{ if demand node } j \text{ self-assigns, meaning a} \\ \text{facility is allocated to site } j \\ 0, \text{ otherwise} \end{cases}$$

The *p*-median model formulation is thus:

$$\textit{Minimise} \;\; Z = \sum_i \sum_j a_i \, d_{ij} \, x_{ij}$$

Subject to
each demand must assign to a facility:

$$\sum_j x_{ij} = 1 \quad \forall i$$

each demand assignment is restricted to what has been located:

$$x_{ij} \le x_{jj} \quad \forall i,j, i \ne j$$

exactly *p*-facilities are located:

$$\sum_j x_{jj} = p$$

and integer restrictions on decision variables:

$$x_{ij} = 0, 1 \quad \forall i,j.$$

The above formulation can be classified as a binary-integer programming problem, and represents the case where each node is both a point of demand and a potential facility site. The formulation can easily be tailored to restrict facilities or demand to specific subsets of nodes. The model determines how each demand node is to be served. The sense of the objective will ensure that each node is served by its closest facility. It is important to recognise that it is not necessary to require that the assignment variables x_{ij}, where $i \ne j$, be integer. This keeps the number of integer variables equal to the number of nodes that are listed as potential facility sites. The above model can be solved directly using a general purpose integer linear programming solution procedure. Because the formulation represents a rather tight structure, an optimal relaxed linear programming (LP) solution often meets the integer constraints. Consequently, a branch and bound process is frequently not required. Unfortunately, the model is quite large in terms of the number of variables and constraints (n^2 variables and $n^2 - n + 1$ constraints). Using a general-purpose maths programming package is, therefore, limited to problems of less than several hundred nodes. This means that it is unrealistic to integrate GIS and general linear programming/integer programming (LP/IP) software to solve *p*-median location models.

Minimising average distance can still leave individual demand points long distances from their closest facility. In the context of emergency facilities, such long distances may be viewed as too far for adequate service. For example, a suburban neighbourhood should be within a mile and a half of a fire station to be considered adequately covered by fire protection services. In planning emergency services, average distance does not capture the urgency of the service, and is commonly replaced by the use of a maximal distance or time standard. A demand is defined as *covered* if a facility has been located within some predefined service standard (e.g. maximal service distance or time). Even though all demand points are to be served, the focus is to provide as many as possible with a level of service which meets some minimally acceptable standard (called 'service coverage'). The development of location models utilising the coverage concept have taken two principal directions: location set covering (Toregas and ReVelle 1972) and maximal covering (Church and ReVelle 1974). Location set covering models minimise the number of facilities needed to cover all demand. The maximal covering location model locates a fixed number of facilities in a manner that coverage is maximised. The maximal covering location model and the location set covering model form the basis of a large class of location models.

Minimising weighted distance and maximising coverage are perhaps two of the most popular objectives that have been developed in location science. It is important to note that a covering model, like maximal covering, can be represented in a completely different model structure from that of a median problem. Consider:

$$y_i^2 = \begin{cases} 1, \text{ if demand }_i \text{ is not covered} \\ 0, \text{ otherwise} \end{cases}$$

$$x_i = \begin{cases} 1, \text{ if a facility is located at site } j \\ 0, \text{ otherwise} \end{cases}$$

$N_i = \{ j \mid d_{ij} \le S \}$, the set of sites which can provide coverage to demand i.

The maximal covering location model formulation:

$$\textit{Minimise } Z = \sum_i a_i y_i$$

Subject to
defining if demand i is covered:

$$\sum_{j \in N_i} x_j + y_i \ge 1 \ \forall \ i,$$

locating exactly p facilities:

$$\sum_j x_j = p;$$

and integer restrictions:

$$x_j = 0, 1 \ \ \forall j \ \text{ and}$$

$$y_i = 0, 1 \ \ \forall i$$

The above formulation is very compact and contains only n+1 constraints and 2n variables. It is structured to minimise the amount of demand not covered. This is mathematically equivalent to maximising the number that can be provided coverage. The maximal covering model is easy to solve optimally for problems involving thousands of nodes using general purpose integer programming software.

There are two important theoretical issues associated with the median and covering models. First, most covering models can be cast as a special form of some type of a median problem (Church and Weaver 1986). This can be done for the maximal covering model as follows:

$$d_{ij} = \begin{cases} 0, \text{ if } d_{ij} \le \mathrm{S} \\ 1, \text{ otherwise} \end{cases}$$

By using the distance values d_{ij} in the p-median problem, the sense of the objective is transformed from minimising weighted distance to minimising the number that are served beyond the distance S. Therefore, any maximal covering problem can be solved using a p-median solution technique by first making a distance transformation (Church and ReVelle 1976; Hillsman 1984). It is, therefore, not surprising to see that the design of the location–allocation module in ARC/INFO, for example, utilises a p-median solution technique and provides the capability to transform a covering model into a median problem format.

The second major issue is that the p-median and maximal covering problems are non-deterministic polynomial (NP)-hard. Practically, this means that specific instances of the p-median and maximal covering problems will be difficult if not impossible to solve optimally in some reasonable amount of computer time. The fact that many vehicle routing, location, and districting problems fall into the class of NP means that applications involving such spatial optimisation problems integrated into GIS must be designed to involve potentially difficult problem instances. Since general purpose solution procedures are only capable of solving median problems that are less than several hundred nodes and since median and covering problems can be represented by thousands of nodes in a GIS, it is virtually impossible to consider solving such problems optimally. There are two approaches that have gained favour in the modelling literature: heuristics and Lagrangian relaxation with limited branch and bound. Lagrangian relaxation with branch and bound has been employed to solve the p-median problem to an exact optimum or to identify a solution within a known percentage of optimality. This technique requires a special purpose code and can be quite sensitive to parameter settings in a specific application. Lagrangian relaxation has been used to solve problems of sizes up to a thousand nodes (Beasley 1993). Heuristics have been developed to solve the *p*-median problem for three principal reasons: (1) to solve problems whose size falls outside the range of an algorithm; (2) to solve problems considerably faster than an algorithm; and (3) to reduce the costs of implementation and application.

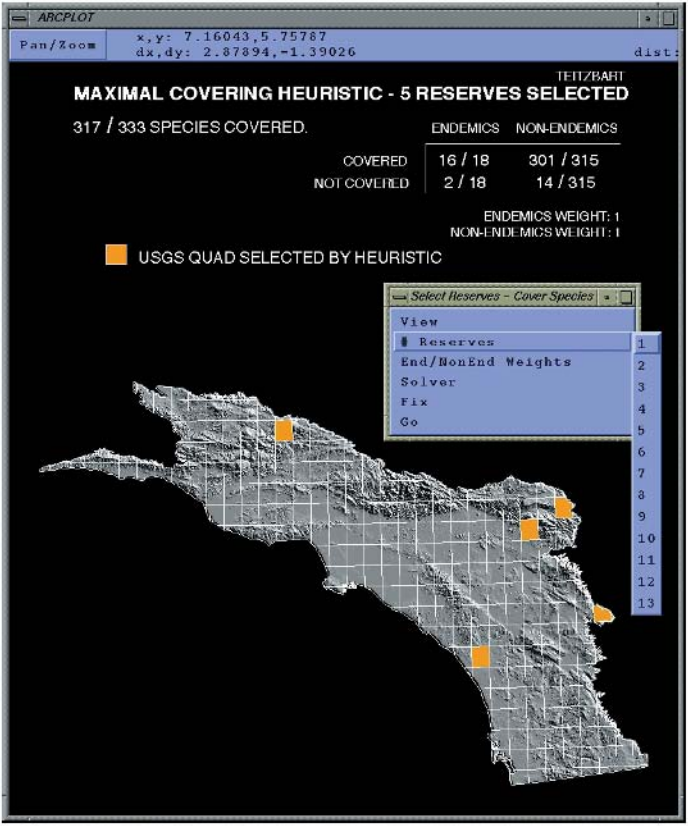
Several types of heuristic have been designed to solve the *p*-median problem, and some of them can be easily implemented into GIS. Such heuristics include: vertex substitution, simulated annealing, greedy adding and dropping, semi-greedy, GRASP, genetic, TABU, and hybrids like GRIA (global regional interchange algorithm: Densham and Rushton 1992). Details of these procedures are provided by Church and Sorensen (1996). The procedure that has received the widest recognition is the simple vertex substitution process of Teitz and Bart (1968). The Teitz and Bart procedure starts with a pattern of *p* facilities. At each step of the heuristic, a candidate site is selected and tested to see if it can be used as a substitute for one of the current facility sites. If any such substitution can be made which yields an improvement in the objective, then a switch (or substitution) is made. The switch that is made is the one which yields the best improvement in weighted distance. In one test, Rosing et al (1979) found that the Teitz and Bart heuristic performed flawlessly, repeatedly identifying optimal solutions with different starting solutions for a 49-node problem. This led the researchers to believe that the process was virtually independent of the starting solution. It is now recognised that even though such a process is relatively robust at finding good, if not optimal, solutions, its performance is controlled by the potential existence of many local optima (Church and Sorensen 1994).

It makes sense to pick a heuristic like Teitz and Bart for solving median and covering models and integrate it into a GIS, based on the premises that it is easy to explain, relatively easy to code, produces very good results and is relatively fast at converging to a final answer. The only problem is that it will not always generate the same final solution unless it is initiated with the same starting solution. Such behaviour is not a common characteristic in GIS, as users expect GIS functions to produce the same answer, unless changes in the data are made. One design approach to this problem is to develop a procedure to determine a starting solution which will always be the same for the same dataset. In this case, the heuristic will always terminate with the same result on the same dataset. This is, in fact, the approach taken by ESRI in developing a location-allocation module for ARC/INFO. A second approach is to acknowledge the limitations of the heuristic and allow the capability of restarting the heuristic with multiple starting configurations, until a relatively simple stopping rule is reached (Church and Sorensen 1994). Doing this can increase the probability of identifying the optimal solution to a given *p*-median problem. Unfortunately, such a process may not always produce the same result.

As stated above, median and covering models have been used for many types of planning problem. An example is given in Plate 14, which depicts the results of a reserve site selection problem for biodiversity protection in southern California. The problem involves the selection of a set of sites which contains as many species as possible. In essence, this is a form of the maximal covering location problem as defined for reserve site selection (Church et al 1996). The result was produced by the location–allocation module of ARC/INFO which was applied to a special logical network of sites and species presence (Gerrard et al 1996). The complete application involves a user interface programmed in the ARC macro language. By customising specific applications, it is possible to aid analysts in data exploration, testing model sensitivity, and visualising results.

6 WHAT'S OPTIMAL WHEN SOLVING A LOCATION MODEL?

Twenty years ago, location problem sizes in geographical analysis were relatively small (i.e. less than 100 nodes, see Hillsman 1984; Swain 1974). One of the principal reasons for this was the difficulty in collecting, storing, and retrieving relevant data. Much of this can now be easily provided by GIS. Instead of having a list of feasible sites, it is now possible to rate them based upon multiple attributes. Many of the classic location–allocation models do not explicitly take advantage of this type of information. Brill (1979) has stated that when a model is solved optimally, that optimal solution may be inferior (i.e. not Pareto optimal) when another objective or criterion is introduced. This means that optimality can be defined for only what has been explicitly included in a model. To support the search for solutions which meet the needs of the decision-maker, it is important to produce close-to-optimal solutions (i.e. solutions which meet high levels of explicitly stated objectives



```
x,y: 7.16043,5.75787
dx,dy: 2.87894,-1.39026
```

MAXIMAL COVERING HEURISTIC - 5 RESERVES SELECTED

317 / 333 SPECIES COVERED.

	ENDEMICS	NON-ENDEMICS
COVERED	16 / 18	301 / 315
NOT COVERED	2 / 18	14 / 315

ENDEMICS WEIGHT: 1
NON-ENDEMICS WEIGHT: 1

🟧 USGS QUAD SELECTED BY HEURISTIC

Plate 14

An example of reserve site selection for biodiversity in southern California.

as well as good performance in terms of those issues that are still a part of the decision-making process but not included in the model. The bottom line is that GIS provides a richer data fabric, upon which more complicated location models need to be defined and solved. Integrated solution techniques need to be capable of searching for high-performing alternatives, instead of just one solution (Lombard and Church 1993). This area of analysis represents a new frontier for GIS and decision support.

7 LOOKING TO THE FUTURE

GIS has had a substantial impact in the field of location model application, in that it has presented a valuable way to organise spatial data for locational search processes. Such searches include retail site location, emergency services location, and factory location. GIS has also allowed many applications to share data easily within a 'corporate' organisation, sometimes creating unanticipated value for location model applications. The integration of multi-facility location models, like *p*-median and maximal covering, has just begun and is generally oriented towards the location of public facilities. Although few systems currently support a multi-facility model (like ARC/INFO), this will undoubtedly change. Features like algorithms to solve capacitated facility location problems are needed to meet the needs of many industries.

The biggest issue facing the GIS community in facility location planning is that most location models and their solution counterparts have been developed and tested for green-field planning problems. That is, the models are typically defined and the algorithms designed to solve the case when all of the facilities are new. Even though all current GIS implementations have the capability to fix existing sites into the solution, they do not have the capability to move existing sites when to do so is relatively superfluous to the solution. That is, sites can only be fixed into a solution or out of a solution. This capability only supports one side of what can be called the brown-field planning problem. In brown-field planning (i.e. adding to, taking away, or transforming an existing configuration) there must be the capability to solve for a new configuration which maintains much of what currently exists and which adds or moves specific facilities to better locations. This can be defined formally for the *p*-median problem using the following notation:

E = $\{ j \mid \text{site } j \text{ currently houses a facility} \}$
R = the number of existing facilities that are to be closed
P = the number of existing facilities
P^i = the number of facilities to be added

We can then define a version of the brown-field median location problem by replacing the *p*-facility constraint $\left(\sum_j x_j = P\right)$ with the following two constraints:

(a) Close R of the current facilities:

$$\sum_{j \in E} x_{jj} = P - R$$

(b) Open P new facilities in addition to the R being closed:

$$\sum_{j \in E} x_{jj} = P$$

The above type of model form is relatively flexible, allowing for some existing sites to be closed and/or new ones to be added. At this time, few if any solution algorithms have been designed or tested to consider such important nuances. The fact that planners usually need to incorporate existing facilities into a new pattern, either whole or in part, can only be approached with this added model construct. Expanding GIS location routines to be capable of supporting this type of functionality should be a high priority.

References

Arentze T A, Borgers W J, Timmermans H J P 1996 Design of a view-based DSS for location planning. *International Journal of Geographical Information Systems* 10: 219–36

Armstrong M P, Densham P J, Lolonis P, Rushton G 1992 Cartographic display to support locational decision making. *Cartography and Geographic Information Systems* 19: 154–64

Beasley J 1993 Lagrangian heuristics for location problems. *European Journal of Operations Research* 65: 383–99

Beaumont J R 1981 Location–allocation problems in the plane: a review of some models. *Socioeconomic Planning Sciences* 15: 217–29

Brandeau M L, Chiu S S 1989 An overview of representative problems in location research. *Management Science* 35: 645–74

Brill E D Jr 1979 The use of optimisation in public sector planning. *Management Science* 25: 413–21

Church R L, Loban S R, Lombard K 1992 An interface for exploring spatial alternatives for a corridor location problem. *Computers and Geosciences* 18: 1095–1105

Church R L, ReVelle C S 1974 The maximal covering location problem. *Papers of the Regional Science Association* 32: 101–18

Church R L, ReVelle C S 1976 Theoretical and computation links between the *p*-median, location set covering, and the maximal covering location problems. *Geographical Analysis* 8: 406–15

Church R L, Sorensen P 1996a *Integrating normative location models into GIS: problems and prospects with the* p-*median model*. Technical Report 94–5. Santa Barbara, NCGIA

Church R L, Stoms D M, Davis F W 1996 Reserve selection as a maximal covering location problem. *Biological Conservation* 76: 105–12

Church R L, Weaver J R 1986 Theoretical links between median and coverage location problems. *Annals of Operations Research* 6: 1–19

Current J R, Schilling D A 1987 Elimination of source A and B errors in *p*-median location problems. *Geographical Analysis* 19: 95–110

Current J R, Schilling D A 1990 Analysis of errors due to demand aggregation in the set covering and maximal covering location problems. *Geographical Analysis* 22: 116–26

Densham P J 1994 Integrating GIS and spatial modelling: visual interactive modelling and location selection. *Geographical Systems* 1: 203–19

Densham P J, Rushton G 1991 *Designing and implementing strategies for solving large location–allocation problems with heuristic methods*. Technical Report 91–10. Santa Barbara, NCGIA

Densham P J, Rushton G 1992 A more efficient heuristic for solving large *p*-median problems. *Papers in Regional Science* 71: 307–29

Dobson J E 1979 A regional screening procedure for land use suitability analysis. *Geographical Review* 69: 224–34

Eiselt H A 1992 Location modelling in practice. *American Journal of Mathematical and Management Sciences* 12: 3–18

Garey M R, Johnson D S 1979 *Computers and intractability: a guide to the theory of NP-completeness*. New York, W H Freeman and Co.

Gerrard R A, Stoms D A, Church R L, Davis F W 1996 Using GIS models for reserve site selection. *Transactions in GIS*

Goodchild M F 1977 An evaluation of lattice solutions to the problem of corridor location. *Environment and Planning A* 9: 727–38

Goodchild M F 1979 The aggregation problem in location–allocation. *Geographical Analysis* 11: 240–55

Hillsman E L 1984 The *p*-median structure as a unified linear model for location–allocation analysis. *Environment and Planning A* 16: 305–18

Hillsman E L, Rhoda R 1978 Errors in measuring distances from populations to service centers. *Annals of Regional Science* 7: 74–88

Huber D L, Church R L 1985 Transmission corridor location modelling. *Journal of Transportation Engineering* 111: 114–30

Idrisi 1996 WEIGHT command. *http://www.idrisi.clarku.edu/PRODUCTS/specindx.htm#DECISION SUPPORT*.

Lombard K, Church R L 1993 The gateway shortest path problem: generating alternative routes for a corridor location problem. *Geographical Systems* 1: 25–45

McHarg I L 1969 *Design with nature*. New York, The Natural History Press

Miller H J 1996 GIS and geometric representation in facility location problems. *International Journal of Geographical Information Systems* 10: 791–816

ReVelle C S 1987 Urban public facility location. In Mills E S (ed) *Handbook of regional and urban economics*, Vol 11. Amsterdam, Elsevier Science: 1053–96

ReVelle C S, Swain R 1970 Central facilities location. *Geographical Analysis* 2: 30–4

Rosing K E, Hillsman E L, Rosing-Vogelaar H 1979 A note comparing optimal and heuristic solutions to the *p*-median problem. *Geographical Analysis* 11: 86–9

Schilling D A, Jayaraman V, Barkhi R 1993 A review of covering problems in facility location. *Location Science* 1: 25–55

Stout J 1989 Peak-load staffing: what's fair for personnel and patients? *Journal of Emergency Medical Systems* 14: 73–4

Swain R 1974 A parametric decomposition approach for the solution of uncapacitated location problems. *Management Science* 21: 189–98

Teitz M B, Bart P 1968 Heuristic methods for estimating the generalised vertex median of a weighted graph. *Operations Research* 16: 953–61

Toregas C, ReVelle C S 1972 Optimal location under time or distance constraints. *Papers of the Regional Science Association* 28: 133–43

Veregin H 1995 Developing and testing of an error propagation model for GIS overlay operations. *International Journal of Geographical Information Systems* 9: 595–619

Wesolowsky G O, Love R F 1971 Location of facilities with rectangular distances among point and area destinations. *Naval Research Logistics Quarterly* 18: 83–90