Chapter Two. Literature Review.

2.1 Strategic Planning, Decision Theory and Decision Support Systems

Decision support systems, which are based on decision theory, have been developed to guide planners, business managers and government officials in making strategic decisions that will steer the future operation of their organization. Although strategic planning initiatives will vary from organization to organization, the planning process, such as planning the geographic distribution of dry hydrants, generally consists of several distinct and successive steps. The first step is defining the goals of the organization or a particular project. These goals may have single or multiple objectives (Eastman 2003). Birkin et al. (1996) point out that objectives generally involve increasing or maximizing effectiveness and efficiency. More specifically, an organization's objectives may be to minimize risk, increase sales, improve customer service and/or to develop better products.

The second general step in the planning process is to identify the options that the organization may have in meeting its objective or objectives (Birkin et al. 1996). Eastman (2003) defines this set of alternatives as the "decision frame" and Birkin et al. (1996) note that frequently, a combination of alternatives or options must be chosen when making a decision. A single criterion or multiple criteria may be used to decide between the available options. Criteria are usually classified as two distinct types: factors and constraints. Factors, which most commonly occur on continuous, quantitative scales, "enhance or detract from the suitability of a specific alternative for the activity [or decision] under consideration" (Eastman 2003, pg 146). In an example in which an organization may be determining where to place a factory (a site suitability model), a factor may be expressed as the desire to locate the factory as close to existing roads as possible to reduce the cost of building additional, factory access roads. On the other hand, constraints serve to "limit the alternatives under consideration" (Eastman 2003, pg 146). Constraints are absolute – they completely eliminate some decision alternatives from the decision frame. In the example of determining a site for a factory, a constraint may be that the factory cannot be located within a kilometer of a wetland to prevent the degradation of these sensitive, unresilient habitats.

The third step in planning initiatives involves choosing one or more of the options identified in the previous step. To make these decisions, "decision rules" are established by developing a systematic, structured process that combines the previously identified criteria or options. In the site suitability scenario described above, the constraint of not placing the factory within a kilometer of a wetland can be combined with the factor that aims to locate the factory close to an existing road to form the decision rule. The results of the decision rule are known as the "decision set", or the decided upon course of action that will best meet the stated objectives. In the case of planning dry hydrants locations, the decided upon hydrants sites would be the decision set.

Problem solving in planning scenarios is not necessarily linear (Birkin et al. 1996). The process will generally flow from identifying the objectives, exploring options and deciding upon how to best meet the objectives. But to a smaller degree, planning initiatives may also be cyclical. That is, while exploring options in meeting the objectives, planners may discover that they overlooked some critical objectives. Similarly, while examining the results of their analysis, they may identify more options that could be useful in meeting the current or future objectives. In other words, elements in the decision set may be useful in future planning activities.

2.2 Dry Hydrant Literature

Literature regarding dry hydrants appears to be rather scarce. But the National Fire Protection Association (NFPA) Document 1142 - "Standard on Water Supplies for Suburban and Rural Fire Fighting" provides many detailed specifications on the installation, and maintenance of dry hydrants. These specifications include the materials used to construct dry hydrants and the steps necessary to install a dry hydrant. Regarding the planning of a dry hydrant location, this document is rather brief does recommend that all affected agencies and private parties should be included in the planning process. Specifically, this document recommends that the following considerations should be included in while planning the location of future hydrants:

1) Current and future population and building trends
2) Property values of the area protected
3) Potential for loss

4) Fire history
5) Current water supply systems
6) Potential for water supply sources – constructed or natural
7) Cost of project
8) Other factors of local concern (NFPA 1999, pg 1142-25).

Appendix C of this document, titled "Water Hauling," discusses methods of estimating fire truck travel speeds, which is imperative in determining the area that could be served by a dry hydrant within a specified maximum allowable response time. The details of these NFPA methods are discussed in Chapter Three of this thesis where they are applied. This thesis addresses the third, fourth, fifth, sixth and eight considerations listed above. However, the three other considerations can be incorporated into GIS-based dry hydrant planning if the necessary data are available and can be georeferenced.

In 1993, the US Forest Service (1993) published the "Dry Hydrant Manual: A Guide to Developing Alternative Water Sources and Delivery Systems For Rural Fire Protection." In this document, they recommend incorporating a dry hydrant plan into a fire department's "master fire plans." Chapter Two of this document, titled "Choosing Hydrant Locations," begins by describing the various water sources that could be used to facilitate a hydrant, which were listed in Chapter One of this thesis. The USFS then encourages future dry hydrant users to enact a "Water Source Survey" in their community. This water source inventory plan encourages local authorities to work with staff and volunteers to compile a list of potential water sources and to record data about each one. Such data include the type of water source (e.g., ponds, streams, existing cisterns and tank), the accessibility of the water source from an existing road and the owner of the property on which the water source is found.

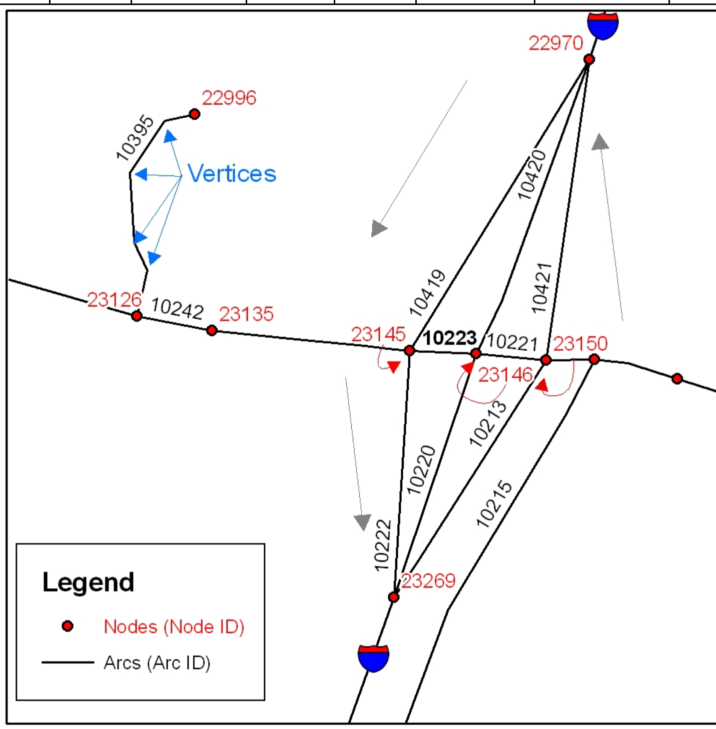
This document then instructs its users to plot the feasible hydrant locations on a county road map, which can then "assist county planners in selecting strategic hydrant sites for water supply" (USFS 1993, pg 2-3). This document then states that "past experience has shown that dry hydrants should be placed at an interval of one in every 3 square miles. This spacing would require tankers to travel approximately 3 miles round trip to any location" (USFS 1993, pg 2-3). This approach problematic and infeasible. First, the spatial configuration of the road network may be irregular and result in the inability to access particular areas within this 3-mile drive. Also, a prominent natural feature, such as a mountain or lake, may prevent fire suppression vehicles from accessing a portion of a particular three square mile track. Furthermore, certain roads may be too narrow or too sinuous to facilitate a fire truck traveling at the needed speed.

Other problems may also be encountered when attempting to plan the location of multiple hydrants. The placement of one hydrant is likely to affect the placement of all future hydrants. For example, the first hydrant installed will only serve a limited area within an acceptable response time. The next hydrant to be installed must compensate for this limited coverage provided by the first hydrant. Then another hydrant must be installed to cover additional unserved areas. The cascading effect of installing hydrants "on the fly" could likely result in an excessive number of hydrants needed to cover a given area. This potentially inefficient method could easily overstretch a fire management agency's finances or could place an undue burden on taxpayers to fund these excessive hydrants. It is problems of this nature that location-allocation models were designed to address.

2.3 Road Network GIS Data Models

Humans, goods and services, such as fire suppression services, typically move through road networks rather than moving in straight lines across landscapes. Hence, modeling road networks is crucial when analyzing transportation operations. A number of network data models have been implemented in computer aided design (CAD) software, MapInfo software and ESRI software. ESRI software includes three network data models, including ArcInfo coverages, ArcView shapefiles and their geodatabase geometric networks. Because the ArcInfo coverage data model is used in this thesis, it will be used to demonstrate the elements of a road network. Furthermore, several of the LAM journal articles reviewed in subsequent portions of this chapter refer to some elements or a road network and as a result, road networks must be discussed before proceeding. A sample of an ArcInfo road network and its associated attribute tables is shown in Figure 2.1 below.

FNODE	TNODE	ARC_ID	FENAME	FETYPE	CFCC	LENGTH_FT	VDOT_SPEED	TO_FROM_MIN	FROM_TO_MIN	FROM_ELEV	TO_ELEV
23250	23269	10213	I-95 Access		A63	1079.1	20	0.613	-1.000	0	0
23146	23269	10220	I-95		A11	1164.8	65	0.204	0.204	0	0
23150	23146	10221	Mudd Tavern	Rd	A31	317.4	45	0.080	0.080	0	1
23269	23145	10222	I-95 Access		A63	1119.3	20	0.636	-1.000	0	0
23146	23145	10223	Mudd Tavern	Rd	A31	301.3	45	0.076	0.076	1	0
23135	23126	10242	Mudd Tavern	Rd	A31	347.8	45	0.088	0.088	0	0
22996	23126	10395			A74	1092.3	5	2.482	2.482	0	0
23145	22970	10419	I-95 Access		A63	1547.2	20	0.879	-1.000	0	0
23146	22970	10420	I-95		A11	1426.4	65	0.249	0.249	0	0
23150	22970	10421	I-95 Access		A63	1374.4	20	-1.000	0.781	0	0



NODE_ID	DEMAND_WEIGHT
22970	2
22996	2
23126	2
23135	2
23145	3
23146	3
23150	3
23269	4

Figure 2.1. An example of a road network data model. The spatial features (arcs and nodes) are shown in the top left, the node attribute tables is show in the bottom right, and the arc attribute table is shown at the top.

Roads are comprised of arcs or links that are joined to one another at intersections. Each link is comprised of vertices that represent a series of georeferenced points with *x* and *y* coordinates. Nodes exist at the dead-end of an arc, at the intersection of two or more arcs, or where the nature of an arc changes, such as a four-lane divided highway condensing down to two lanes. Both arcs and nodes have attribute tables that contain data describing each of these individual features. Each arc's attribute table contains a reference to the nodes found at its end points, one being the "from node" (FNODE) and the other being the "to node" (TNODE), depending on the direction in which the arc was digitized and its intersection with another road. Hence, intersecting arcs can be identified through these tabular entries rather than by a spatial search for two arcs intersecting at a common location in continuous space.

Arcs representing a road segment typically have attributes including its name and road type. In the US Census Bureau's Topologically Integrated Geographic Encoding and Referencing (TIGER) system road data model, this road type is represented by the census feature class code (CFCC) attribute field. A speed limit can be assigned to each of these CFCC classes and as a result, the length of time required to traverse each link can be established using the equation *distance = rate * time*. This time specification is one of the many impedances that can be assigned to arc to indicate the "cost" of traversing it. Different impedance values can be assigned to each direction of travel. For example, it may require two minutes of time to travel from the FNODE to the TNODE because it is downhill. But it may require three minutes to travel from the "to node" to the "from node" because that direction of travel requires an uphill climb. The "FROM_TO_MIN" field and the "TO_FROM_MIN" field in Figure 2.1 exemplify this idea of network impedance costs. One-way streets can also be modeled using this concept of impedance by specifying a negative value in the impedance field, such as Arcs representing the on and off interstate ramps in Figure 2.1 (Arc_ID 10222, 10419, 10421, and 10213). Overpasses and underpasses can also be modeled in a GIS road network. Doing so is accomplished by assigning an elevation value to an arc's end-point and beginning-point; if the elevation values of two intersecting arcs do not match, then the system is directed to not permit the movement from one arc to the other. In Figure 2.1, a turn from the overpass arc #10221 to the interstate highway arc #10220 would not be permitted due to the unmatching elevation values in the arc attribute table.

2.4 Location-Allocation Models

2.4.1 Overview of Location-Allocation Models

A wide variety of location-allocation models (LAMs) have been under development for decades and the literature that discusses these models spans many disciplines including geography, planning, operations research, business management, systems engineering and economics. Scott (1970) provides a detailed summary of LAM research conducted prior to 1970 and cites literature dating as far back as Valinski's 1955 article titled "A Determination of the Optimum Location of Fire-Fighting units in New York City". LAMs continue to be refined and developed today and an entire journal, *Studies in Locational Analysis*, is dedicated to this subject as well as other related topics.

Scott defines location-allocation problems as geographically oriented problems that aim to place optimally a number of central supply centers in a plane with established, dispersed locations that have demand for goods or services being distributed from those supply centers. The second component of these problems typically assigns flow between the limited supply centers and the demand locations they are to serve.

2.4.2. The Transportation Problem

Some simple location-allocation problems are composed only one of the above components, such as the problem that Scott labels as "The Transportation Problem." This problem addresses the situation in which demand *and* supply locations are already established and merely aims to assign the flow of goods and services from these established centers to the demand locations in a manner that will minimize the overall transportation costs. The solution to this problem is given by minimizing the equation:

$$\text{Minimize} \quad Z = \sum_{i-1}^{m} \sum_{j=1}^{n} d_{ij} x_{ij} \tag{1}$$

subject to:

$$\sum_{J=1}^{m} x_{ij} \le q_i \qquad (i = 1, 2, \dots, m) \tag{2}$$

$$\sum_{i=1}^{m} x_{ij} \ge r_j \qquad (j = 1, 2, \dots, n) \tag{3}$$

where:

n = the number of demand locations

m = the number of central facilities or supply locations

d_{ij} = the distance between i and j

x_{ij} = the magnitude of some flow between i and j

r_j = the requirements (demand) of the j^{th} destination point or region

2.4.3 The Weberian Location Problem

On the other hand, the "Weberian Location Problem" is purely locational and only addresses the first component discussed above. It uses calculus techniques to place a *single* supply location in a planar, continuous coordinate space, the so-called and well-known "median point" or "point of minimum aggregate travel." Love et al. (1988) refer to this type of LAM as "Site-generating location-allocation models" because it establishes a new location in continuous space, as opposed to selecting a location from user-defined set of candidate locations. Solving this problem involves minimizing the Pythagorean distances between the supply center and the demand locations, as given below:

$$\text{Minimize } Z = \sum_{i=1}^{n} r_j [(X^* - x_i)^2 + (Y^* - y_i)^2]^{1/2} \qquad (4)$$

where:

X^*, Y^* = X and Y coordinates of the central facility to be located

x_i, y_i = the x and y coordinates of the demand locations

(see previous section for all other notations)

Iterative calculus methods are then employed to solve for X^* and Y^* - the coordinate location for the central facility to be established (Scott 1970).

2.4.4 Optimal Partitioning Problem

When the number of supply centers to be establish is greater than one, solutions become more complex. A problem of this type where there is no upper bound on the supply capacity of the supply center can be classified into the category that Scott (1970) calls the "Optimal Partitioning Problem". When facilities are to be placed in continuous space, the location of the supply centers and the assignment of demand locations to these centers can be determined by minimizing the following equation, which builds upon equation 2 above:

$$\text{Minimize } Z = \sum_{i=1}^{m} \sum_{i=1}^{n} \lambda_{ij} r_j [(X^* - x_i)^2 + (Y^* - y_i)^2]^{1/2} \qquad (5)$$

subject to:

$$\sum_{i=1}^{m} \lambda_{ij} = 1 \qquad (j = 1, 2, \ldots, n) \tag{6}$$

$$\lambda_{ij} = \left\langle \begin{matrix} 1 \\ 0 \end{matrix} \right. \tag{7}$$

where:

λ_{ij} = a zero-one variable such that whenever $\lambda_{ij} = 1$, then activity *ij* is brought into existence at some positive level; and otherwise whenever $\lambda_{ij} = 0$

The λ_{ij} variable specifies whether or not the j^{th} demand point is assigned to the i^{th} supply center. The restriction imposed in equation 6 ensures that each demand location (*j*) is assigned to one and only one supply center (*i*). Again, iterative calculus techniques are then employed to solve the equations for X* and Y*, the coordinate locations of the supply centers to be established.

When the Optimal Partitioning Problem is applied in discrete space where the distances between potential or candidate supply locations and demand points are pre-calculated, such predetermined distances through a road network, the problem can be solved by minimizing the following equation, which builds upon equation 1 above.

$$\text{Minimize } Z = \sum_{i=1}^{m} \sum_{j=1}^{n} d_{ij} r_j \lambda_{ij} \tag{8}$$

subject to:

$$\sum_{i=j}^{n} \lambda_{ij} = 1 \tag{9}$$

$$\sum_{i=1}^{n} \lambda_{ii} = m \tag{10}$$

$$\lambda_{ii} - \lambda_{ij} \geq 0 \qquad \begin{matrix} (i = 1,2, \ldots, n) \\ (j = 1,2, \ldots, n) \end{matrix} \tag{11}$$

$$(i \neq j)$$

$$\lambda_{ij} \geq 0 \tag{12}$$

Equation 9 denotes that $\lambda_{ij}=1$ when demand point *j* is assigned to supply center *i*. Equation 10 stipulates that only *m* supply facilities will be located. Equations 11 and 12 require that if a supply center is established at the same location as a demand point, then that demand point will be assigned to its coincidental supply location. Love et al. (1988) classify this type of problem as a "site selecting location-allocation model" because a set of candidate locations in discrete space is given from the onset of the problem.

2.4.5 The P-median Problem

A majority of the LAMs commonly used today are based on the models discussed above. These fundamental models are then modified to meet certain objectives, a few of which will be discussed below. The p-median problem aims to place a fixed, user-specified number of supply centers (*p*) in a manner that minimizes the total distance from the supply centers to the demand locations (ESRI 1994). This model is typically used in the private sector and the classic example that demonstrates its intended use is the problem of placing warehouses. A commodities distributor must transport his or her products from warehouses to existing retail outlets. To minimize shipping and operational costs, the warehouses storing the goods should be located in a configuration that minimizes the *aggregate* travel distance between each retail outlet and the nearest warehouse. Similarly, an airline may also use a p-median model to establish a new hub in a manner that maximizes operational efficiency.

2.4.6 The Attendance Maximizing LAM

While the p-median LAM aims to minimize the total distance traveled from supply centers to demand locations, the "attendance maximizing" LAM seeks to place facilities in areas of higher demand density (ESRI 1994). The premise in this model is that the likelihood of attendance (i.e., using a facility) is function of distance – people are more likely use facilities that are closer and if the distance is too great, then people are likely to not use the facility at all. Hence, this model can be attractive to those planning the locations of libraries, parks or stores. The ArcInfo implementation of this model allows the user to specify the distance at which attendance will not occur.

2.4.7 The Covering Set LAM

The covering set location model aims to determine the minimum number of facilities needed to "cover" an entire population within a maximum allowable response time or distance. It second objective is to then determine where to locate these facilities. For example, a county's health care system may aspire to ensure that the entire population is within an eight-minute travel time from the nearest rescue station. Because constructing and manning these rescue stations will require substantial and variable funding, the county would also prefer to minimize the number of these stations. The covering set model would be an appropriate approach to meeting these objectives. Mathematically, the covering set problem can be written as (Owen and Daskin 1998):

$$\text{Minimize} \quad \sum_j c_j X_j \qquad (13)$$

$$\text{Subject to:} \quad \sum_{j \in N_i} X_j \geq 1 \qquad \forall i \qquad (14)$$

$$X_j \in \{0,1\} \qquad \forall j \qquad (15)$$

Where:

C_j = fixed cost of siting a facility at node j

S = maximum acceptable service distance or time

N_i = set of facility sites j within acceptable distance of node i

The objective function (13) is to minimize the overall cost of establishing the facilities and therefore, the variable c_j is included. The overall cost of building eight high-cost facilities may be higher than building nine low-cost facilities. The first constraint (14) ensures that every demand location is served by at least on facility. Unlike the maximal covering location problem (discussed below), the degree of demand at each node is considered to be perfectly homogeneous. If the demand for the facility is scattered throughout an area of interest and the maximum allowable response time is small, then a large number of facilities will likely be required.

2.4.8 The Maximal Covering Location Problem

Church and ReVelle (1974) identified the potential for cost-ineffective planning when using the covering set approach. For example, if six fire stations could serve 80% of the population, six additional fire stations may be required to serve the remaining 20% of the population, as shown in Figure 2.2 below.

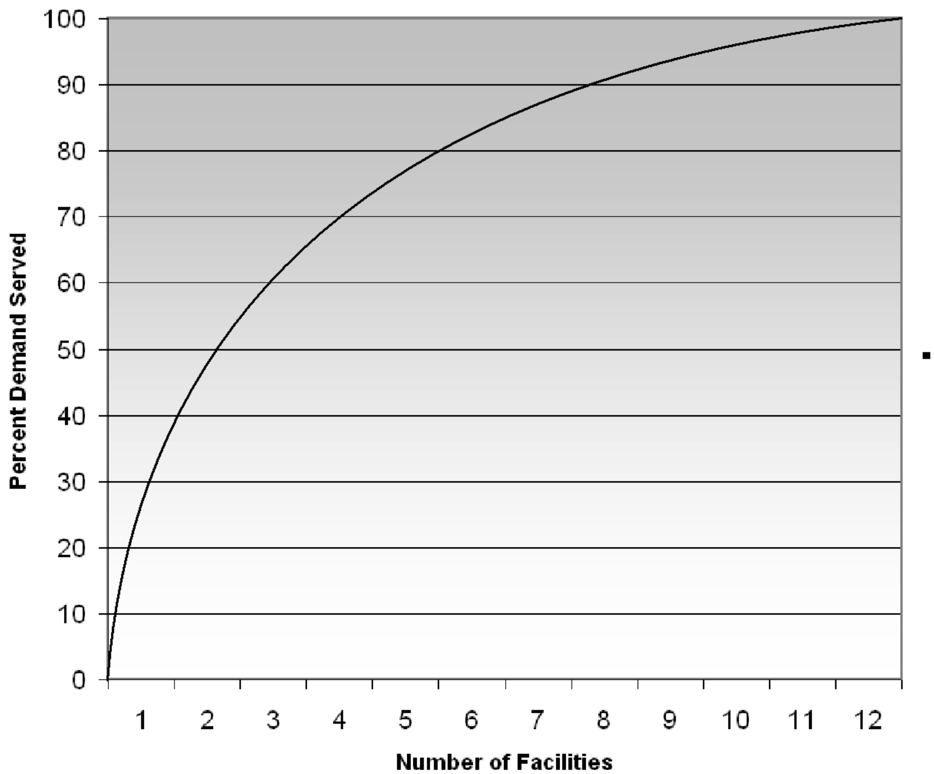


Figure 2. 2. This curve demonstrates the trade-off between the number of facilities that are established and the percentage of the population that can be served within a maximum response distance or time.

In situations in which the cost of establishing one of these facilities is high, operation budgets can quickly be over extended and exhausted. To address this reality, they proposed a location model that aims to "cover", within a specified response distance or response time, the maximum number of demand locations with a fixed number of facilities: the maximal covering location problem (MCLP).

ESRI (1994) mathematically represents the MCLP as:

$$\text{Max } Z = \sum_{i=1}^{n} w_i z_i \qquad (16)$$

Subject to:

$$\sum_{i=1}^{m} y_j = p \qquad (17)$$

$$\sum_{i \in N_i}^{m} y_j \geq z_i, \forall i \qquad (18)$$

$$y_j = 0,1, \forall j \qquad (19)$$

$$z_i = 0,1, \forall i \qquad (20)$$

where:

- i = demand location
- j = candidate facility location
- n = number of demand locations
- Ni = $\{j \mid d_{ij} <= S\}$ the set of facilities j that can reach node i within the maximal service distance/time S
- m = number of candidate facility locations
- p = the number of facilities to locate
- w_i = the weight at demand node i
- d_{ij} = the shortest distance from node i to node j
- y_j = 1 if a facility is located at site j, otherwise 0
- z_i = 1 if demand node i is covered, otherwise 0

Equation (17) restricts the number of established facilities to p, a user-defined parameter. Equation (18) stipulates that node i is not 'covered' unless one or more facilities at site j are in the set of Ni facilities that can reach node i within distance S (the maximum service distance or time). Lines (19) and (20) give the facility location and coverage decision variables, respectively. The first variable in line (16), the weight at each demand location, can be specified or manipulated by the user to emphasize or deemphasize areas that have more or less need for a particular service. Most LAMs, including the ArcInfo implementation of LAMs, are performed on road networks and the representation of supply and demand for a good or service is consequently spatially restricted to these nodes. If few nodes exist in the areas of high population density, the population weight can then be increased at these nodes to more accurately represent the true population distribution in the road network data model.

Based on the MCLP model, as the number of supply facilities (*p*) increases, the amount of additional demand served by each additional facility decreases. In economic terms, the marginal benefit of each additional facility decreases. When plotted on a graph (Figure 2.2), this relationship allows the users of MCLP models to visually examine the cost effectiveness of increasing the number of facilities. In addition to varying *p*, users of MCLP models can also manipulate the maximum acceptable response time or distance (*S*). As *S* increases, the amount of demand that can be reached from each facility will increase and therefore the number of facilities needed to reach the demand will likely decrease. Hence, both P and S have a positive relationship with the amount of demand served by the configuration of facilities.

If the Maximal Covering Location Problem LAM is run multiple times and the number of facilities is increased during each subsequent run, the MCLP LAM can serve as a surrogate for the Covering Set Problem. This approach is exemplified in Figure 2.2, where the 12^{th} facility covers the entire population. However, this approach of determining the covering set is computationally rigorous. In Figure 2.2, the MCLP LAM had to be executed 12 times to determine the covering set, rather than running the Covering Set Location Problem only once.

2.4.9 The MCLP with Mandatory Closeness Constraints

In the same article, Church and Revelle (1974) also introduce the maximal covering location problem with mandatory closeness constraints. This variation of the MCLP allows planners who aim to cost-effectively serve their population to specify two levels of maximal response distance or response time. The first distance is the *desired* distance (*S*) and the second is the mandatory response distance beyond which no members of the population should be located (*T*). Mathematically, this model is very similar to the MCLP model, with one additional constraint:

$$\sum_{i \in M_i}^{m} y_j \geq 1, \forall i \qquad (21)$$

where:

$M_i = \{j | d_{ij} <= T\}$ or the set of facilities j that can reach node i within the mandatory distance T. Since $T > S$, Mi contains the set N_i

This variation does not guarantee that the entire population will be covered, particularly if P, S and/or T are small, as shown below.

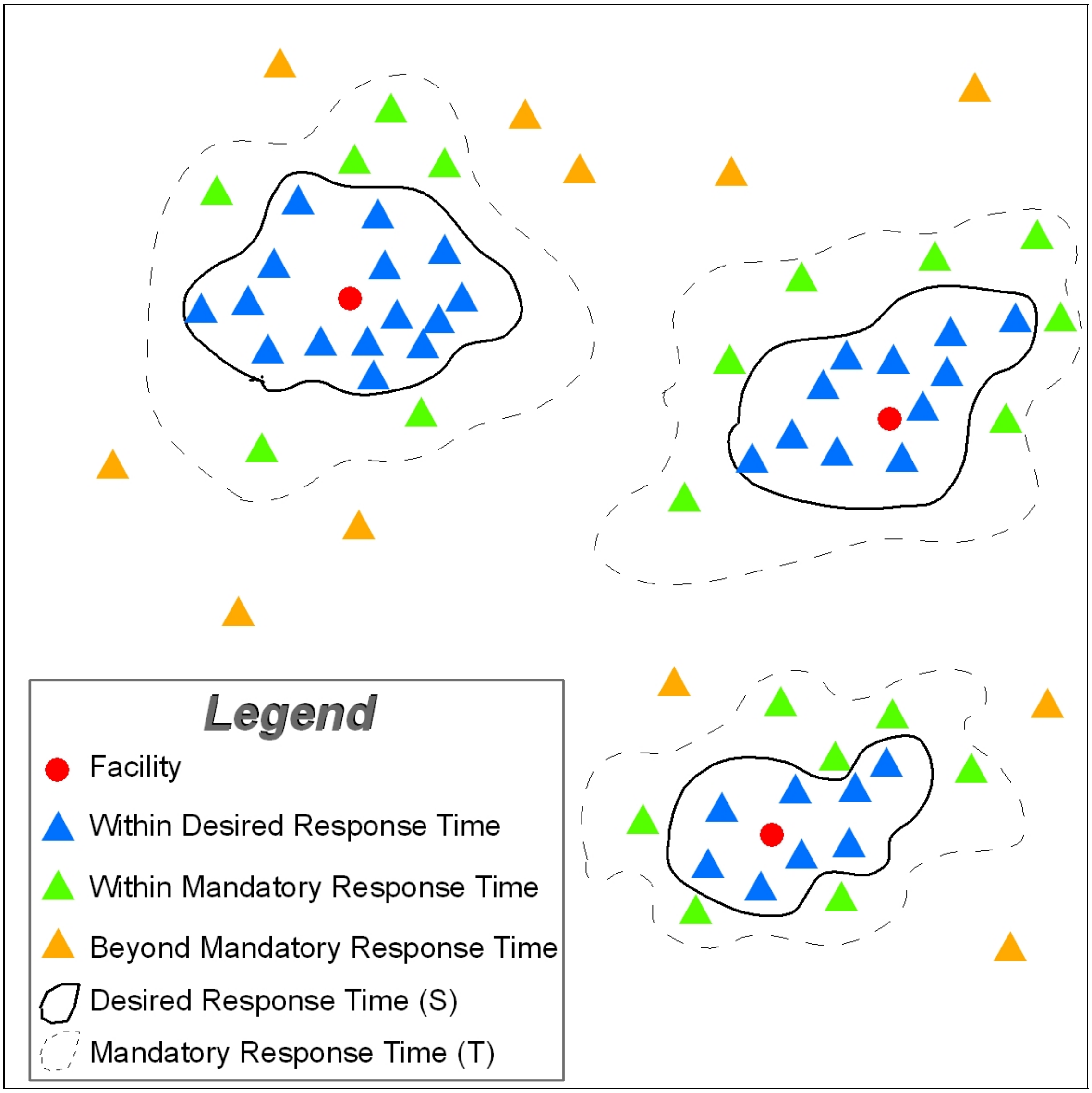


Figure 2.3. An optimal solution to the MCLP with Mandatory Closeness Constraints with three facilities. A majority of the demand locations are within the desired response zone, fewer are in the mandatory response zone, and even fewer are beyond the mandatory response zone.

2.4.10 ArcInfo Implementation of LAMs.

The Network module ArcInfo software has implemented six LAMs from which its users can choose. These include the P-median Problem, the Attendance Maximizing Problem, the MCLP and the MCLP with mandatory closeness constraints. The Covering-set Location Problem LAM is not included in this software suite, but an iterative use of the MCLP can be used as a surrogate for the Covering- set Location Problem.

Although some features of ArcInfo Network module allow its users to assign demand to arcs, those using the LAMs can only distribute the demand to nodes. However, LAM users can assign *costs* to the arcs. For example, the minutes required to traverse a particular arc can be assigned to each arc. Hence, the maximum allowable response distance in the MCLP problem can be expressed as a maximum allowable response *time*.

2.4.11 Heuristic Procedures in Location-Allocation Models

Location-allocation models have the potential to require an enormous number of calculations. First, under most circumstances, the shortest distance or travel time between each candidate site and each demand location must be calculated. If 150 demand locations can be served by 75 candidate sites, then 11,250 of these distances must be calculated, a very reasonable number of calculations for modern personal computers. But the process of optimally selecting central facility sites from a large candidate pool has the potential to require an enormous number of hypothetical combinations. The number of possible combinations for facility sites can be represented as (ESRI 1994):

$$\frac{n!}{p!(n-p)!}$$

where:

n = the number of candidate locations

p = the number of sites to choose

In the context of placing dry hydrants in Spotsylvania County, Virginia, 639 candidate locations exist. If only 15 hydrants are to be established in the county, then 3.0463 x 10^{35} comparisons would have to be made to optimally choose the best configuration, requiring a tremendous amount of computing time. Obviously, this time requirement is unacceptable and Scott (1970) remarks that most location-allocation problems are solvable only through the use of heuristic procedures, which have been developed to address problems of this nature. Heuristic methods are those that consist of "a sequence of trials yielding approximate results with control of the progression towards an acceptable final result" (Webster's Dictionary of Computer Terms, 1994). Hence, heuristic procedures cannot guarantee an optimal solution, but if well designed, they tend to converge in the direction of an optimal solution.

Scott (1970) summarizes a heuristic procedure proposed by Cooper in 1963. This procedure makes an initial "guess" at a solution in situations in which the number of facilities to be established is a user-defined parameter. In the MCLP, this initial guess may be to select the candidate sites that reach a large number of demand nodes. Each of these demand sites will then be assigned to the nearest facility. The procedure then swaps some of these initial sites with other candidate sites. If the swap results in a better solution, then the swap is stored as a part of the solution. If one of these swaps results in drastic reduction in the desired goal, then the site examined in that iteration is flagged as very uneconomical and is permanently deleted from the list of candidate locations. Hence, the number of subsequent comparisons that must be made has been reduced. This process continues until no other combinations are likely to produce a better result.

Densham and Rushton (1991) comment on the well-designed heuristic developed by Teitz and Bart in the late 1960s to solve location-allocation problems. They note that their colleagues' heuristic is usually accurate (i.e., it does in fact converge on the optimal solution), fast and versatile. But Densham and Rushton identified some opportunities to improve this heuristic in a manner that requires fewer computations and less use of computer memory. One of these improvements concerns the number of demand nodes that are assessed when a substitution for a particular candidate site is being evaluated. Teitz and Bart's approach reallocates every demand location to the new, temporary set of candidate sites. On the other hand Densham and Rushton's procedure, known as the global-regional interchange algorithm (GRIA), only reallocates the demand nodes that are affected by the swap (those that were originally allocated to candidate site being swapped out and those that are within the potential service range, what Sorensen and Church (1995) call the "cut-off distance", of the site that is swapped in). Such refinements of the heuristics used in location-allocation models have allowed for these complex problems to be solved in a microcomputer environment. Indeed, the GRIA heuristic is the default heuristic used in ArcInfo software's LOCATEALLOCATE command.

Sorenson and Church (1995) discuss the potential for decreased quality in the solution of location-allocation problems, particularly the P-median problem, when using the GRIA heuristic. If the truly optimal solution contains a distance between a central facility and a demand node that exceeds the cut-off distance, then the short-cut resulting from the used of the GRIA heuristic would prohibit the identification of the optimal solution. Perhaps this is the reason why the ArcInfo software gives its users the option of using the GRIA heuristic, the default, or the Teitz and Bart heuristic.

2.5 Other approaches to Emergency Response Assessment and Planning

2.5.1 Emergency Response Assessment and Planning Before the GIS Was Commonly Used

GIS has been used extensively in planning the location of emergency facilities, including fire and rescue stations, hospitals or other medical care centers, and even mobile EMS units. This section of this chapter will review a handful of the many articles that speak to the use of GIS in emergency planning. However, to provide contrast two articles predating the widespread use of GIS will be discussed first.

In 1968, Jane Hogg worked with fire officials in the Bristol County Borough of the United Kingdom in planning the location of fire stations both within the borough and in the surrounding areas as well. This area had six existing fire stations and the presiding fire chief identified thirteen sites that could host future stations. The objectives of her research strongly mirror the objectives that most planners face in similar initiatives: the number of fire stations to establish and where to locate them. Hogg had originally intended to establish her criteria based on the monetary loss of property resulting from a fire. With this as her criteria, she stated that the optimum number of stations would be the "number beyond which any additional station would not be paying for itself." But because it was impossible to incorporate monetary values into her model (the reasons for this were not elaborated upon), she reorganized her criteria to be the "minimization of the total time spent by all appliances journeying to fires." Thus, the nature of her problem strongly mirrors the p-median problem that is briefly elaborated upon in the section of this chapter that discusses location-allocation models.

To assess demand for fire stations, she partitioned the study area into square areas that measured one kilometer by one kilometer – a square tessellation. The number of fires occurring in each square was summed, resulting in a raster-like density map of the county. Adjacent squares that had similar fire densities were homogenized to form 15 larger sub-areas of the county. Hogg then established a "centre of gravity" point within each sub-zone by "reading off the median appliance journey [to historical fires] along the x and y axis of the national grid. The junction of the roads nearest to this point (x,y) was then treated as the centre of gravity." The result of this tactic reduced the demand for fire services across an entire sub region, the largest of which was approximately 49 km^2, to a single point. Hogg didn't elaborate on her reasons for doing so. Perhaps the limited computing resources of her time compelled her to simplify her problem and then reduce the number of calculations required to execute her model.

Hogg estimated travel times from each candidate location to each centre of gravity by studying historical response times and riding along with the fire department on actual calls. Hence, she established a matrix of travel times (t_{ij}) from each station to each sub-region centre of gravity. Her problem was compounded by the fact that each appliance (i.e., fire truck) has a limited capacity in meeting demand. So, if a sub-region required five trucks, but the station within that sub-region only housed three trucks, then two trucks from nearby stations would be needed. Her model can then be expressed as

$$\sum_{i=1}^{n}\sum_{j=1}^{m} t_{ij} d_{ij}$$

where d_{ij} = the number of historic journeys made from station i to sub-area j. Her objective was then to minimize this function.

The results of Hogg's model showed, as expected, that the aggregate travel time between each station and each demand point decreased exponentially as the number of stations increased. In other words, when a large number of stations was theoretically added to the county, the reduction in the aggregate travel time was small.

Due to the high operational costs of urban fire suppression, most of which is used to pay fire-fighting and administrative staff, Plane and Hendrick (1977) worked with the City of Denver, Colorado to analyze the existing spatial configuration of the city's fire stations in an effort to reduce costs while not compromising the quality or effectiveness of fire suppression services. As the city expanded geographically, new fire stations were constructed to provide sufficient fire-fighting "coverage" to these newer, localized areas. The long-term effect of establishing additional fire stations "on demand" was what was perceived to be an inefficient spatial configuration of fire stations covering the entire city. The goal was then to determine if some existing fire stations should be relocated and/or closed.

To establish demand for fire stations, the city fire officials identified the location of fire hazards and then ranked them into five ordinal classes (the criteria for establish a fire hazard was not specified in this paper, neither were the criteria for ranking them). Next, the city fire officials assigned a maximum acceptable response time for each hazard class with the higher hazards having a lower acceptable response time. Using historical response times from fire stations to the scene of a fire, regression equations were derived to predict the response time as a function of distance.

The researchers then applied a hierarchical "covering-set" location model. Covering set models aim to determine the number of facilities (fire stations in this case) needed to "cover" the entire population of a given area within a maximum allowable response time. Once they determined the number of facilities needed to cover the city, they then executed a secondary model (the hierarchy) to maximize the use of existing stations and thereby avoid the cost of constructing a new fire station or relocating any existing fire stations.

After running the hierarchical set-covering model, the fire officials recognized some counter intuitive results that were later attributed to the short response time demanded by a handful of high hazard structures on the periphery of the city. After carefully considering these hazardous outlying structures, the officials decided that their stringently applied response times could safely be relaxed. The hierarchical model was then re-executed. The results of this separate run led to the elimination of two existing fire stations, which, according to this model, did not reduce the quality of fire suppression services to its citizens and businesses.

2.5.2 Approaches to Emergency Response Assessment and Planning Using GIS

Walsh et al. (1997) used road network modeling in ArcInfo to assess the service areas of hospitals in a 16 county area surrounding Charlotte, NC. They then simulated the effects of doubling the demand for hospital care on the service area of each hospital. This study took advantage of nearly all of the features available in ArcInfo's networking allocation algorithms: supply, demand and impedance. Supply was assigned to each node representing an existing hospital based on the number of available beds and attending physicians. They established demand by utilizing existing hospital discharge records that were geo-coded to the 5-digit zip code of each patient. Finally, they established impedance along the network links or arcs by estimating the travel time needed to traverse each link based on the estimated speed for each road type.

The authors then allocated each arc to the nearest hospital. If the capacity of a particular hospital was exhausted, then the arcs were assigned to the next closest hospital with remaining capacity. Most of the service area generated from the allocated arcs were contiguous and the hospitals were generally located near the center. But in rural areas with smaller hospitals, each hospital's capacity was quickly exhausted and consequently the "un-served" population had to travel to larger capacity hospitals which in some cases required them to traverse two counties to access health care. Walsh et al. (1997) then simulated a twofold increase in health care demand while holding each hospital's capacity constant. Consequently, each hospital's service area areally shrank, which thereby created areas that were un-served by any hospital in the entire region.

Two researchers at the University of Nevada, Las Vegas generated service areas for the region's existing fire stations to determine "un-served" routes that will likely be used to transport radioactive waste to the Yucca Mountain waste facility that is currently under construction (Parentela and Sathisan 1995). Using census data, they also mapped the population distributions that may possibly be affected by an accident involving a waste transport. The Federal Highway Transportation Administration recommends a 10 minute response time to hazardous waste spills and these researchers determined that an alarming amount of the Las Vegas area's periphery is beyond this 10 minute response range. This research strongly implied that hazardous materials should be rerouted or that additional hazardous waste response units should be established.

Estochen et al. (1998) used service areas generated by GIS network analysis to test the effectiveness of pre-deploying EMS ambulances to area in and surrounding Des Moines, Iowa that have historically experienced a high frequency of traffic accidents. Using a spatial database, they mapped traffic accidents occurring over a five year period and then determined the percentage of the accidents that were located within 5, 7 and 10 minute intervals from existing, stationary EMS stations. They then manually moved some ambulances to areas of higher accident densities (the "pre-deployment") and they found that the *average* response times to accident locations only decreased by 24 seconds. However, the percentage of accidents within a 5 minute threshold increased from 56.4% to 63.2% for morning rush hour accident and from 58.7% to 65.8% for afternoon rush hour accidents. This pre-deployment only reduced 7 minute response times by 2%. Accidents with a 10 minute threshold showed very little change, but overall, response capabilities were improved by the hypothetical pre-deployment.

2.6. Model Selection in Dry Hydrant Planning

As discussed, location-allocation models excel in situations that require facilities to be established from a large candidate pool of potential sites. In particular, the Maximal Covering Location Problem LAM is designed to aid emergency planners in determining where to place emergency response facilities so they can maximize the demand served within a specified maximum response time. When run iteratively, the MCLP can provide planners with a means to evaluate cost-effectiveness of installing additional facilities and the number of facilities needed to meet all of the demand in a particular area, as shown in Figure 2.2. The ArcInfo implementation of this model allows the travel time of fire suppression vehicles to be modeled on a road network. Hence, the MCLP model is used in this thesis.