Netw Spat Econ (2007) 7:129–152
DOI 10.1007/s11067-006-9003-6

Location of Alternative-Fuel Stations Using the Flow-Refueling Location Model and Dispersion of Candidate Sites on Arcs

Michael Kuby · Seow Lim

Published online: 19 December 2006

Abstract The Flow Refueling Location Model (FRLM) is a flow-intercepting model that locates *p* stations on a network to maximize the refueling of origin–destination flows. Because of the limited driving range of vehicles, network vertices do not constitute a finite dominating set. This paper extends the FRLM by adding candidate sites along arcs using three methods. The first identifies arc segments where a single facility could refuel a path that would otherwise require two facilities at vertices to refuel it. The other methods use the Added-Node Dispersion Problem (ANDP) to disperse candidate sites along arcs by minimax and maximin methods. While none of the methods generate a finite dominating set, results show that adding ANDP sites produces better solutions than mid-path segments or vertices only.

Keywords Continuous location · Link · Flow intercepting · Hydrogen refueling · Finite dominating set · Vehicle range

The environmental, geopolitical, and financial costs of the world's dependence on gasoline-powered vehicles are becoming increasingly obvious. Although hybrid vehicles increase the efficiency of vehicles, they may not be the long-term answer. Many people in government, industry, academia, NGOs, and the general population are looking to alternative-fuel vehicles powered by electricity, biofuels, natural gas, or hydrogen. The transition to alternative-fuel vehicles, however, will not be easy. New infrastructure will be needed to produce, transport, store, and distribute alternative fuels. California is developing a statewide network of hydrogen refueling stations, and similar initiatives are being considered in Florida, the upper Midwest, and the Southwest (California EPA 2005). The National Renewable Energy Lab (NREL) examined the possible pathways to transition to a national network of hydrogen refueling stations using GIS (Melendez and Milbrandt 2005). Locating refueling stations to serve the maximum demand possible will be a key to the early success of efforts such as these.

M. Kuby (✉) · S. Lim
School of Geographical Sciences, Arizona State University, Tempe, AZ 85287-0104, USA
e-mail: mikekuby@asu.edu

Kuby and Lim (2005) developed the Flow-Refueling Location Model (FRLM) to aid in the optimal location of a network of refueling stations for alternative-fuel vehicles. The FRLM extends the flow-capturing model developed by Hodgson (1990) and also by Berman et al. (1992), who referred to it as "flow-intercepting." Flow-capturing models, which are structurally similar to the maximum-cover model problem (Church and ReVelle 1974), locate a given number of facilities to maximize the traffic flow volume that passes by the facilities. The concept of flow capturing provides an excellent foundation for locating refueling stations, although one major modification is needed. For refueling vehicles with a limited driving range, the assumption that a flow can be captured by a single facility located anywhere along the path may not hold. Short trips may be refuelable by a single facility, but longer trips will require a combination of facilities spaced adequately along the shortest path.

The problem of optimally locating facilities on transportation networks is often much easier if the infinite number of potential locations along the arcs of a network can be reduced to a finite dominating set of discrete points (Hooker et al. 1991). For instance, Church and Meadows (1979) developed a method for adding a set of network-interception points (NIPS) to the set of vertices to create a finite dominating set for the set-cover model. For the flow-capturing model, the nodes of the network constitute a finite dominating set to which the search can be limited (Hodgson 1990; Berman et al. 1992). Their result, however, does not carry over to the FRLM. Kuby and Lim (2005), and Kuby et al. (2004), showed that certain paths may not be refuelable by a single facility located at the nodes on the path, but could be refueled by a facility located mid-arc. The search for optimal locations of refueling stations, therefore, cannot be limited to the nodes of a network.

We have not identified a finite dominating set for the FRLM. Instead, this paper compares three methods for adding candidate sites on the arcs. We first develop a method to identify segments of arcs in which a single facility would be able to refuel a path for which two or more facilities at vertices would otherwise have been required, and use these "mid-path segments" to augment the set of vertices. The other two methods use the Added-Node Dispersion Problem (ANDP) to disperse candidate sites uniformly along the arcs using minimax and maximin methods. Developed by Kuby et al. (2005), it was previously shown to improve solutions to the p-dispersion problem. Here we show that augmenting the set of vertices using the generic ANDP method substantially improves the FRLM's solution and generally outperforms our custom method for generating additional candidate sites specifically for the FRLM. The remainder of this paper formally introduces the FRLM (Section 1), the mid-path segment method (Section 2), and the minimax and maximin versions of the Added Node Dispersion Problem (Section 3). We then introduce a test network (Section 4) and compare the locations of the candidate sites generated by the three methods (Section 5). We then test these sets of added candidate sites against each other and against the original vertices of the network for solving the FRLM for various numbers of facilities and vehicle ranges (Section 6) before concluding and suggesting directions for future research (Section 7).

1 The flow-refueling location model and other fuel-station location models

In the flow-refueling location model, demand consists of origin–destination (O–D) pairs, rather than nodes, to be covered. Optimal location of fueling stations has been treated prior to Kuby and Lim (2005), but not from the perspective of serving origin–destination demands. Goodchild and Noronha (1987) maximized market share based on both node-to-facility trips and traffic counts on network arcs. Bapna et al. (2002) dealt specifically with infrastructure development for alternative fuels. Their multiobjective approach, called the Maximum Covering/Shortest Spanning Subgraph Problem (MC3SP), incorporated vehicle range by adding and/or upgrading enough stations on arc *ij* so that a vehicle of standard range can traverse the arc using alternative fuels. The number of stations needed per arc is determined exogenously based on maximum spacing and the location of existing gas stations. Nicholas et al. (2004) developed a GIS model to locate hydrogen-refueling stations in Sacramento County, California, similar to a p-median model solved by a greedy algorithm adding two stations at a time. Melaina (2003) did not address exact optimal locations, but proposed several methods for estimating the number of initial hydrogen refueling stations that might be required to provide a large proportion of the U.S. population with convenient access.

In contrast to these approaches, the FRLM locates p facilities so as to maximize the path-flow volume that can be successfully refueled. The formulation of the FRLM is based on that of the flow-capturing model, which mirrors that of the maximum-covering location problem (Church and ReVelle 1974). The flow-capturing model departs from the max-cover model by substituting the path index q for the node index i. The FRLM departs from the flow-capturing model by using *combinations* of facilities, rather than single facilities, to cover or refuel a path.

In the flow-refueling problem, the limited range of vehicles implies that certain paths may only be refueled by a combination of adequately spaced facilities (Kuby and Lim 2005). The major innovation of the flow-refueling model was the ability to allow combinations of facilities to successfully refuel vehicles on each path. The set of combinations of nodes capable of refueling a given path can be determined exogenously for each path, based on the vehicle range and the spacing of nodes along the path. This is essentially similar to the generation of covering sets for each node in a set-cover or max-cover model, except that (1) the coverage set may consist of single nodes or combinations of nodes, and (2) we use incidence matrix notation, following Daskin (1995), instead of cover set notation. For consistency, if a single node is capable of refueling a path all by itself, it is considered a combination of one.

$$\text{Max } Z = \sum_{q \in Q} f_q y_q \tag{1}$$

Subject to:

$$\sum_{h \in H} b_{qh} v_h \geq y_q \quad \forall q \in Q \tag{2}$$

$$x_k \geq v_h \quad \forall h \in H; k | a_{hk} = 1 \tag{3}$$

$$\sum_{k \in K} x_k = p \tag{4}$$

$$x_k \in \{0, 1\} \qquad \forall k \tag{5}$$

$$v_h \leq 1 \quad \forall h \tag{6}$$

$$y_q \leq 1 \quad \forall q \tag{7}$$

$$v_h, y_q \geq 0 \quad \forall h, q \tag{8}$$

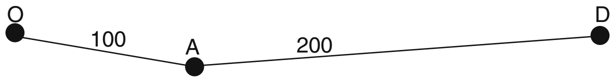
where:

q	Index of O–D pairs (and, by implication, the shortest paths for each pair)
Q	Set of all O–D pairs
f_q	Flow volume on the shortest path between O–D pair q
y_q	1 if f_q is captured 0 otherwise
k	Potential facility location
K	Set of all potential facility locations
h	Index of combinations of facilities
H	Set of all potential facility combinations
b_{qh}	1 if facility combination h can refuel O–D pair q 0 otherwise
v_h	1 if all facilities in combination h are open 0 otherwise
a_{hk}	1 if facility k is in combination h 0 otherwise
x_k	1 if a facility is located at k 0 otherwise
p	The number of facilities to be located

The objective function 1 maximizes the total flow that can be refueled. Constraints 2 require at least one eligible combination of facilities h be open for paths to be refueled. One such constraint is written for each path q. The eligible combinations, which are determined exogenously, may consist of single facilities that can refuel the entire path, pairs of facilities, or more. Constraints 3 hold v_h to zero unless all the facilities in combination h are open. Constraint 4 requires exactly p facilities to be built. The facility location variables x_k are defined as binary variables in 5. Although v_h and y_q are defined in the definitions as 0–1 variables, they can be relaxed in a mixed-integer program as continuous variables with an upper bound of 1, as in 6–8, and still yield an all-integer solution, as explained in Kuby and Lim (2005).

The parameters b_{qh} are key inputs to the FRLM, because they indicate which facility combinations are able to refuel path q for the shortest paths between each origin and each destination in the network. Kuby and Lim (2005) presented an algorithm for determining the set of combinations h and the values of b_{qh}. The algorithm assumes without loss of generality that to refuel a path, a set of facilities must be able to refuel a round trip that is made repeatedly. It assumes that vehicles start the trip at the origin with half a tank of fuel, and that fuel consumption is equal in both directions. For a facility combination h to refuel path q, a range-limited vehicle must be able to reach the first station on the path using a half tank of fuel, and then, refueling as necessary, go to the destination, turn around, and return to the first station. If it can do so, it can refuel there again and make it back to the origin with at least a half tank remaining. These assumptions would allow vehicles to then make the same trip again, or start a different trip with at least half a tank of fuel. These round-trip assumptions also ensure that one-way trips would also conclude with at least a half tank of fuel.

Fig. 1 Sample path from Origin O to Destination D



The basic principles behind this algorithm can be illustrated with the following simple example of a 300-mile shortest path through the network starting from an origin node O, through node A, to a destination node D (Fig. 1). The importance of considering the entire round trip can be illustrated by a single facility located at node O. A vehicle starting at node O could travel to D and return to O if the range is at least 600 miles.

The second principle is that, unlike in the flow-capturing model, the nodes of the network do not constitute a finite dominating set for network location. If, for example, the range of the vehicle is 300 miles, none of the 3 nodes along this path can refuel the round trip without the vehicle running out of fuel. Even with a station at A, it is 400 miles from A to D and back. There is, however, a point at the exact midpoint of the path from which it is exactly 300 miles to D and back and 300 miles to O and back.

The third principle is that, for any vehicle range less than half of the round-trip distance, it is impossible to refuel the trip with a single facility located anywhere on the path. For such cases, a combination of two or more facilities along the path is needed to prevent vehicles from running out of fuel. However, not just any combination of two or more facilities will suffice. For instance, for a range of 200, nodes A and D could successfully refuel the trip, but nodes O and A could not.

The algorithm for determining the combinations of nodes that can refuel a path contains six basic steps (Kuby and Lim 2005):

1: Generate the shortest path for all O–D pairs q, and establish an empty master list of all combinations h.
2: Generate a temporary list of all combinations h of nodes on path q.
3: Remove facility combinations that cannot refuel a vehicle of the given range on path q.
4: If any combination h contains a subset that can also refuel path q, remove it from the combinations list for path q.
5: For all combinations h still on the list for path q:

   Add it to the master list of combinations if it is not already there.
   Record b_{qh}=1 if facility combination h can refuel path q and 0 otherwise.
   Record a_{hk}=1 if facility k is in combination h, and 0 otherwise.

6: Repeat Steps 2–5 for all paths q.

Recall that the network vertices are a finite dominating set for the flow-capturing problem (Hodgson 1990), but not for the flow-refueling problem (Kuby and Lim 2005). Because we have found neither a finite dominating set for the FRLM nor a way to solve the continuous network location version of it, we propose three practical methods for improving the solutions to the FRLM by adding candidate sites along arcs.

2 Candidate site generation: mid-path segments

Figure 1 can motivate the idea behind the first method for providing additional candidate sites. As explained above, if the vehicle range were exactly 300 miles, there is a point in the exact middle of the path, 50 miles east of Node A, where a single facility could successfully refuel the path. If the range of the vehicle were slightly longer, say 350 miles, there would be a segment extending to 25 miles on either side of the midpoint within which a single facility could refuel the route. Our first method for adding candidate sites along network arcs generates segments such as this for every path, and then reduces the number of these segments through a column-reduction process similar to Toregas and ReVelle (1973).

In the following algorithm, d_q is the one-way shortest-path distance for O–D pair q, and r is the vehicle range. The shortest path can be treated as a straight line from Origin O to Destination D as in Fig. 2. The Start and End points define the outer limits of a segment in which a single facility could be located and be able to refuel the entire path all by itself.

Next, we derive the lengths of the segments shown in Fig. 2. The Start of the segment is the point at which the round-trip distance from Start to D and back equals exactly r. Therefore, the distance from Start to D equals $r/2$ and from O to Start equals $d_q-r/2$. By virtue of symmetry, the distance from End to D must also equal $d_q-r/2$. By addition, the distances from O to Start plus from End to D equals $2d_q-r$. Therefore, the distance from Start to End equals $d_q-(2d_q-r)$ or $r-d_q$. Finally, by symmetry again, the distances from Start to Midpoint and from End to Midpoint both equal $(r-d_q)/2$.

The algorithm for generating the complete list of mid-path segments and the paths they cover is as follows.

2.1 Mid-path segment algorithm

Step 1: Generate Mid-path Segments

For every path q:

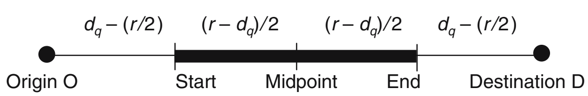
- If $d_q>r$, no single point can refuel path q. Go to next path.
- If $2d_q \leq r$, every point on path q can refuel it, including the origin, destination, and any other vertices along the way. There is no need to add any segments. Go to next path.
- If $d_q \leq r < 2d_q$, generate a mid-path segment that starts at $(d_q-r/2)$ from the Origin, and ends at $r/2$ from the Destination. If the mid-path segment does not contain any vertices, add it to the temporary set of TempSegments. A vertex within any segment will be at least as good as any other point in the segment and quite possibly better because it may also be able to refuel paths that begin, end, or cross at the vertex using other arcs to which it is connected.

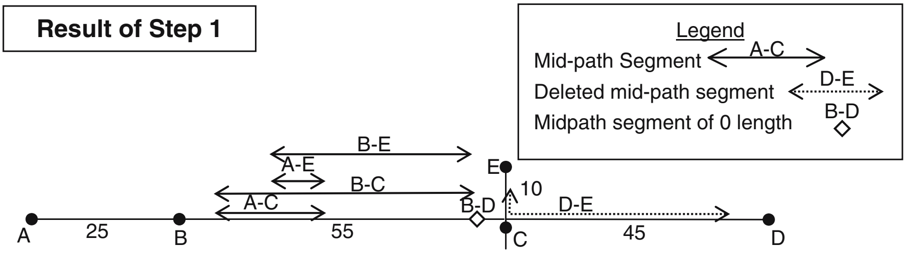
Step 2: Break Up Overlapped Segments into Nonoverlapping Sub-segments

For every arc a:

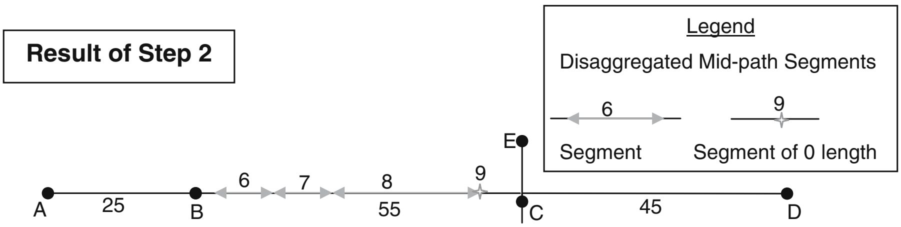
- If arc a contains no segments in TempSegments, go to next arc.
- If arc a contains only 1 segment, or no overlapping segments, add these segments to the GlobalSegment list, keep track of the paths they can refuel in the Paths-Segments list, and go to the next arc.
- If arc a contains overlapping segments, break each uniquely overlapped section into its own nonoverlapping sub-segment. Add the new sub-segments to GlobalSegment list, keeping track of the paths that they can refuel in the

Fig. 2 Mid-path segment definition





Path q	One-way Length	Temp Segment	Location of Segment on Path
A-B	25	None needed—too short	
A-C	80	1	40±10
A-D	125	None possible—too long	
A-E	90	2	45±5
B-C	55	3	27.5±22.5
B-D	100	4	50±0
B-E	65	5	32.5±17.5
C-D	45	None needed—too short	
C-E	10	None needed—too short	
D-E	55	Deleted—includes Node C	27.5±22.5



Result of Step 3

Global-Segment	Paths Refueled	Dominated by	Step 4 Decision
6	A-C,B-C	7	Deleted
7	A-C, B-C, A-E, B-E		Kept
8	B-C,B-E	9	Deleted
9	B-C, B-D, B-E		Kept

Fig. 3 Example of mid-path segment algorithm for a vehicle range of 100

Paths-Segments list. Remove the original segments from the TempSegment and Path-Segment lists. Go to the next arc.

Step 3: Eliminate Dominated Segments

For every segment in Path-Segments:

- Check whether the set of paths it can refuel is a subset of any other segment's set of paths. If so, eliminate this segment. Go to the next segment.

2.1.1 End of mid-path segment algorithm

The mid-path segment algorithm is illustrated in Fig. 3 for a simple 5-node network for a vehicle range of 100. Step 1 generates mid-path segments for six of the ten possible paths. One of these segments consists of a single point in the exact middle of path B–D, which is exactly as long as the vehicle range. The segment for the path D–E is then deleted because it contains node C, which dominates any other point on the segment. Step 2 then examines the overlaps among the five temporary segments and identifies four sub-segments with no overlaps. Step 3 then deletes segments 6 and 8 because segments 7 and 9 dominate them, respectively.

Once the set of segments is finalized, the candidate site can be assumed to be at the midpoint of the segment, but in fact could be anywhere in the segment. Each of these segment midpoints is treated as a candidate site k for the FRLM. These sites are added to the set of vertices, and the combined set of candidate sites is run through the combinations algorithm summarized at the end of Section 1. Given that each segment is capable of refueling at least one path q on its own, these k will be included in the set of combinations h as combinations consisting of a single facility, and b_{qh} will be set to 1 for all the paths each can refuel. They can also be included in combinations of two or more k by the combinations algorithm. A limitation of this method, however, is that the potential usage of these segments *in combination with* other sites is not considered when generating them initially.

3 Candidate site generation: added node dispersion

The second method for adding candidate sites along arcs is the Added Node Dispersion Problem (ANDP), developed by Kuby, Lim and Upchurch (2005). The ANDP disperses p additional discrete candidate sites along the m arcs of a network. Kuby, Lim, and Upchurch developed minimax and maximin versions of the model. To explain these two models, it is necessary to first introduce some terminology.

The term subarc is used to describe the subdivisions of the original arcs of the network. When a new node is added to an arc, the original arc is considered to be divided into two subarcs, s=0 and s=1. If S nodes are added to an arc, it is divided into s=0, 1,..., S subarcs. An arc with no added nodes is considered to have a single subarc (s=0) equal to the entire original arc.

The minimax ANDP minimizes the maximum subarc length in the network. The maximin ANDP maximizes the minimum subarc length. The minimax ANDP adds nodes to arcs so that the longest subarc is made as small as possible. It is motivated by the idea of preventing any long stretches of network without any candidate sites. The maximin ANDP, on the other hand, adds nodes to arcs so that the shortest subarc is made as long as possible. It is motivated by the idea of not wasting any candidate sites by placing them too close to each other.

Like most maximin and minimax objectives, however, these minimax and maximin objectives are characterized by poorly behaved alternate optima. Take the maximin model, for example. If one of the original arcs of the network is very short, it really does not matter where the new nodes are added as long as they do not create a new subarc shorter than the shortest existing arc. To prevent these kinds of poorly behaved alternate optima, Kuby, Lim, and Upchurch included a secondary MinSumMax objective in the models to select the bestdispersed alternate optima. It minimizes the sum of the longest subarcs on each of the original arcs. They proved that equal spacing of added nodes along each original arc is optimal to the MinSumMax objective. They also developed mixed-integer programming (MIP) models to characterize the two ANDP models.

Both minimax and maximin ANDP models can be solved quickly, easily, and optimally with a simple greedy algorithm (Kuby et al. 2005). The algorithm for solving the minimax ANDP identifies the longest subarc at each iteration and adds a node to its original arc—possibly an *additional* node, if one or more have already been added to it. The original arc is then subdivided into one more equal fractions. If the longest subarc is an entire original arc, meaning no nodes have been added to it yet, the algorithm would add a node at the halfway point and split it into two arcs. If the longest subarc is on an original arc that had previously been split in half (or thirds), one would split the original arc into thirds (or fourths), and so on until p nodes had been added to the network.

To maximize the minimum separation, only one small change is required. Instead of adding a node to the original arc with the longest subarc, choose the original arc that will have the longest subarcs *after its next equally spaced subdivision*. Add nodes to arcs one at a time until p nodes have been added.

The difference between the two algorithms can be seen in Fig. 4. For the minimax ANDP algorithm, the longest subarc is reduced from the original 18 units to 11, then 10, and then 9 as the first three nodes are added to each original arc. At $p=4$, a second node is added to the original arc of length 18, splitting it into 3 subarcs of length 6. For the maximin ANDP, the first step is the same, resulting in a maximized minimum length of 9. If the maximin algorithm proceeded the same as the minimax, it would add a node to the second-longest original arc, creating two subarcs of length 5.5 each. If, instead, a second node is added to the arc of length 18, the minimum separation is 6—slightly larger than the minimum of 5.5 if the second node were added to the arc of length 11. This example illustrates the need for the maximin algorithm to choose at each iteration the arc that will have the longest subarcs after a node is added, rather than the arc that currently has the longest subarc.

Kuby et al. (2005) used the minimax and maximin ANDP to generate extra candidate sites along the arcs of a network for solving the p-dispersion problem with a mixed-integer program (Kuby 1987; Erkut 1990). They found that the ANDP candidate sites improved the objective function value of the p-dispersion problem by 22% (for 5 added nodes) to 55% (for 20 added nodes) compared with the vertices of the network, and that there was little difference in the performance of the maximin sites vs. the minimax sites. Section 6 of the present paper will examine if a similar boost in performance can be obtained in the FRLM.

While the mid-path segment and ANDP methods are offered as practical ways to improve the FRLM's solution, some of the candidate sites generated automatically by these methods may not be in practical locations. For instance, a candidate site could be generated on a limited-access highway between exits, where no station could be built. This can largely be prevented in both methods by careful definition of the network. Junction nodes can be placed at all exit points and defined as candidate sites without defining them as trip origin/destination points. With shorter arcs, it is unlikely that the ANDP method will need to place a node in the middle of the arc. The ANDP will be much more likely to add nodes to long uninterrupted arcs representing smaller highways on which a station could be located anywhere. If there is an arc on a limited-access highway that is long enough for the ANDP to add a node in the middle of it, then a user could manually delete it or recommend that it be considered as a new exit or service station location. For the mid-path segment method, recall that Step 1 eliminates

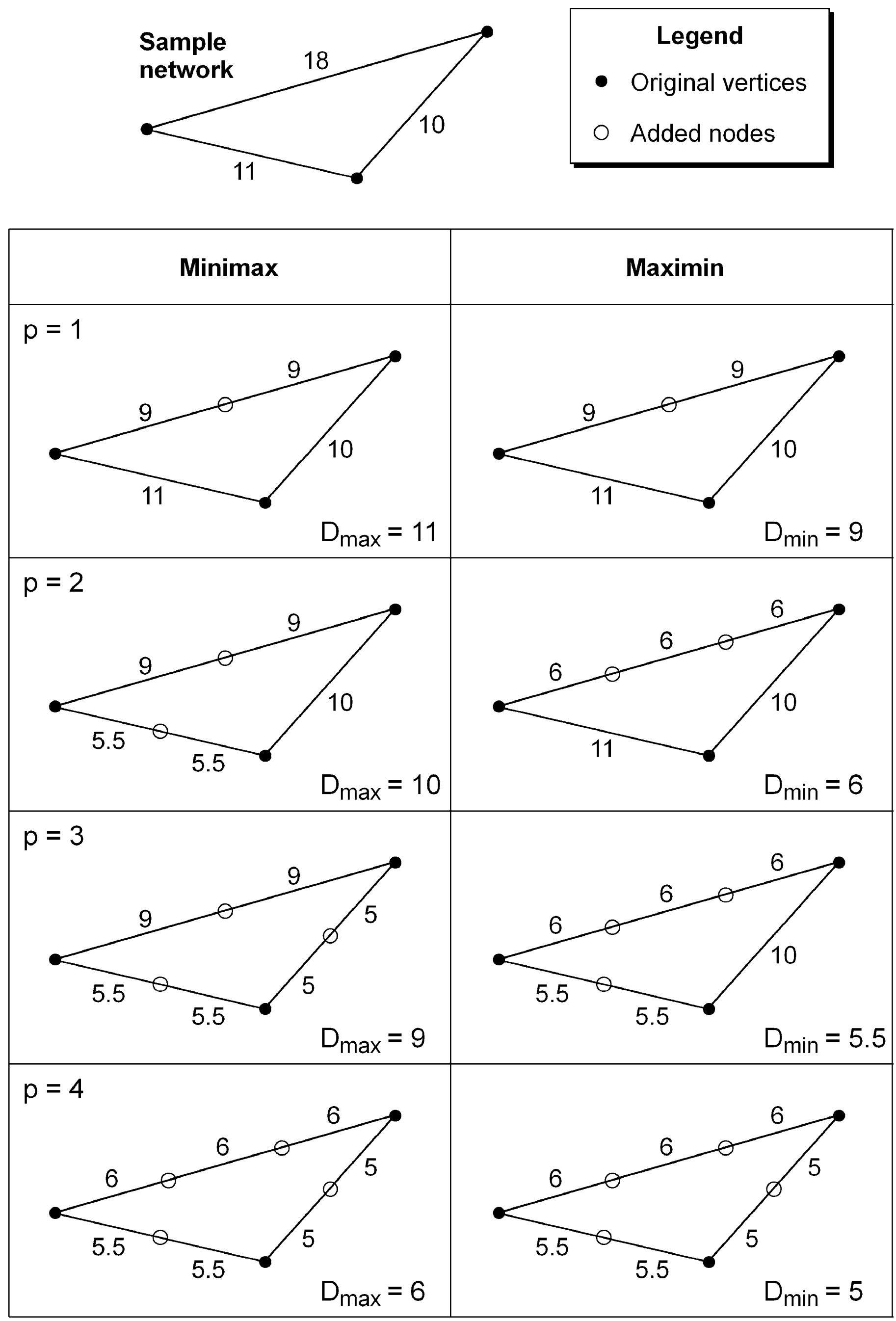


Fig. 4 Minimax and maximin ANDP algorithm results compared for $p=1$ to 4 on a simple network. Reprinted with permission of *Geographical Analysis*

any segments that contain any vertices, and so adding vertices at exits would reduce the size of the problem. Both methods could also be modified easily to exclude from consideration a subset of arcs representing limited-access highways. Neither method is likely to ever add a segment or candidate site in an urban area where arcs are short.

4 Case study and solution procedures

The FRLM was tested on the 25-node sample network used by Kuby and Lim (2005) and Kuby et al. (2005), and before that by Hodgson (1990) and Berman and Simchi-Levi (1988) (Fig. 5). Following Hodgson, the flow volumes f_q in the 25×25 O–D matrix were estimated using a gravity model. Flows were then assigned to their shortest paths.

The FRLM mixed-integer programming model was generated using the Mosel modeling language and solved with the XPress-MP 12.0 solver on a Dell Precision 340 with 2.53 GHz Pentium 4 processor and 1024 MB RAM. The problem size varies with the vehicle range: the longer the range, the more combinations of facilities. With a range of four and no added candidate sites besides the 25 vertices, the model has 300 continuous fractional flow coverage variables y_q, 75 continuous fractional variables v_h for each combination of facilities, and 25 integer x_k variables. The model also has 300 flow-refueling constraints 6, 223 combination constraints 7, and one p-facility constraint 8. With a vehicle range of 12 and no added candidate sites, the number of combination variables goes up to 412, and the number of combination constraints rises to 1274.

Initially, we analyzed a 3×4 matrix of scenarios. We compared four methods for adding nodes along arcs (vertices only; mid-path segments; maximin ANDP; and minimax ANDP) for three different vehicle ranges (four, six, and eight). The mid-path segment algorithm determines the number of added candidate sites. For ranges of four, six, and eight, the number of mid-path segments was 22, 19, and 17 respectively. For consistency, we generated the same number of candidate sites for each range using the maximin and minimax ANDP models. All in all, there are ten different sets of candidate sites used in the 12 scenarios (see Table 1). For each of the 12 scenarios, we solved the model for p=1 to 25 facilities.

5 Comparison of candidate site locations

Before we look at how these additional candidate sites improve the solution of the FRLM, let us first examine where the ANDP and mid-path segment algorithms tend to locate candidate sites. Figure 6(a) compares the 22 segments generated for a vehicle range of four with the same number of ANDP maximin and minimax locations. Figure 6b does the same for the 17 segments generated for a range of eight.

Given a vehicle range of four, the mid-path segment algorithm identified segments on the shorter (but not the shortest) arcs, whereas the ANDP methods located on longer arcs. The midpath algorithm located segments one unit long on all arcs of length three. Any point on these segments can refuel the one-arc paths between the two endpoints. It also generated a segment consisting of a single point at the exact midpoint of each arc of length four. No segments are generated on arcs of length two because the vertices at the end can refuel the one-arc paths, and no segments are generated on arcs of length five or longer because no single point on the arcs can refuel the round trip. In addition, no segments are generated on

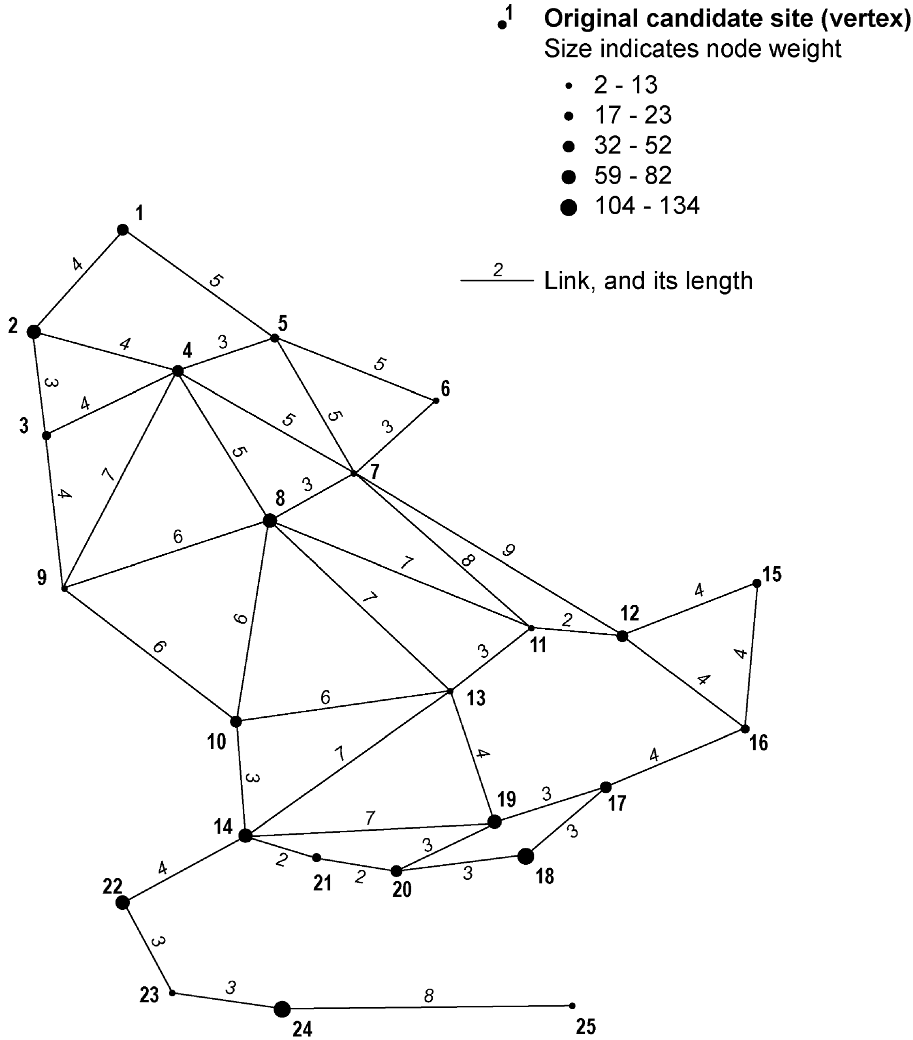


Fig. 5 The test network, based on Hodgson (1990), originally from Berman and Simchi-Levi (1988). Reprinted with permission of *Geographical Analysis*

Table 1 Candidate sites used in each scenario

Vehicle range	Vertices only	Vertices + mid-path segments	Vertices + ANDP maximin sites	Vertices + ANDP minimax sites
4	25 vertices	25 vertices + 22 mid-path segments	25 vertices + 22 ANDP maximin sites	25 vertices + 22 ANDP minimax sites
6	25 vertices	25 vertices + 19 mid-path segments	25 vertices + 19 ANDP maximin sites	25 vertices + 19 ANDP minimax sites
8	25 vertices	25 vertices + 17 mid-path segments	25 vertices + 17 ANDP maximin sites	25 vertices + 17 ANDP minimax sites

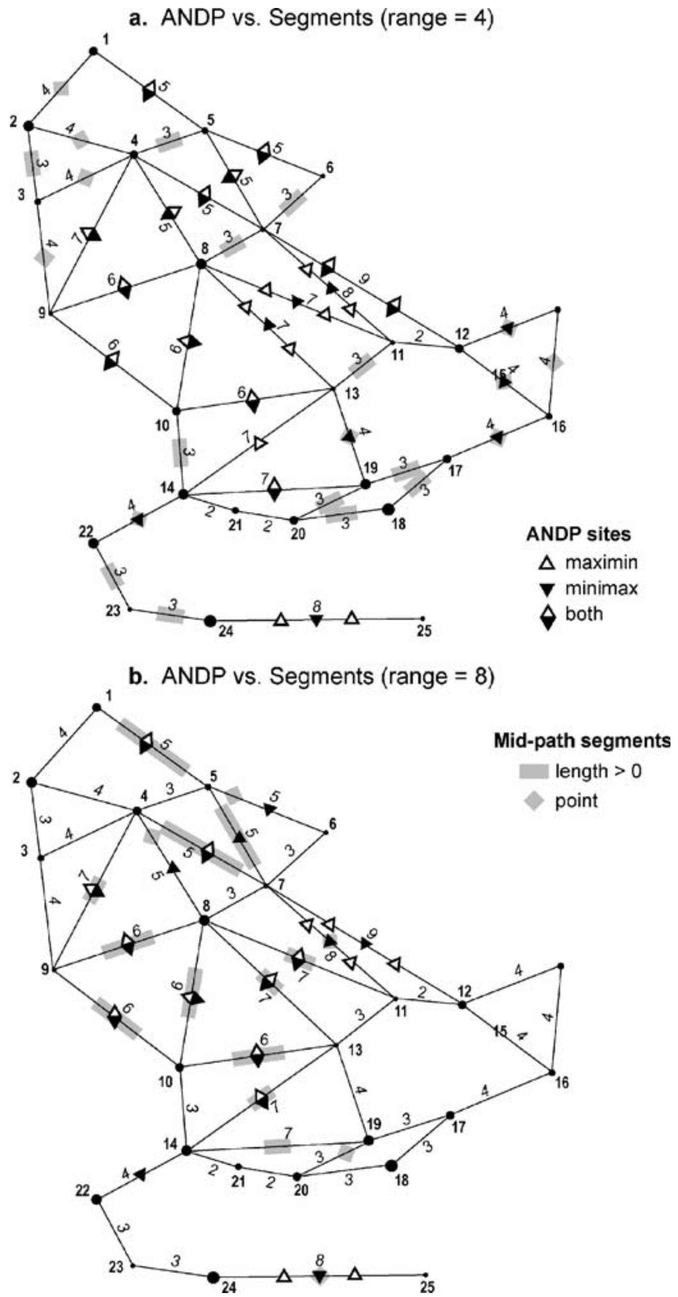


Fig. 6 Comparison of locations of added mid-path segments and ANDP maximin and minimax sites, for a vehicle range of (**a**) four and (**b**) eight. The number of mid-path segments is determined endogenously by the segment algorithm, while the number of ANDP sites is set equal to the number of segments identified

multi-arc paths because the only paths short enough (path 14–21–20) would have had a vertex within the segment.

In comparison, the maximin ANDP added sites only on arcs of length five or more, with no sites added to the same arcs as the segment method. The minimax ANDP also favored longer arcs, but with some overlap with the segment method. The maximin method proceeded thusly: it added a node on all arcs of length nine through six; then a second node to arcs of length nine and eight; then nodes to the arcs of length five; and was partway through adding a second node to the arcs of length seven when the node limit was reached. The minimum separation for any added node at that point was 2.33, found on those arcs of length seven to which two nodes were added.[1] In contrast, the minimax ANDP added single nodes to arcs of length nine through five before adding a second node to the arc of length nine. It then added nodes to four arcs of length four when the node limit was reached. The maximum separation at that point was 4.0 on those arcs of length four to which a node had not yet been added, as well as on the arcs of length eight to which one node had been added.

Unlike the ANDP method, the mid-path segment algorithm depends on the vehicle range. Thus, in Fig. 6(b), based on a vehicle range of eight, the segment algorithm finds segments on arcs of lengths five to eight that are able to refuel the round trips between the arc endpoints, whereas the endpoints themselves cannot. The segments are longest on the one-arc paths of length five, and get gradually shorter until they shrink to a point for one-arc paths of length eight. The round trips on the one-arc paths of lengths three and four no longer require any mid-path segment. A few segments are generated for multi-arc paths, such as the segment on arc 4–8 that can refuel the paths 4–8 and 5–4–8 by itself, and the segment on arc 19–20 that can refuel the path 17–19–20–21 (and others) by itself.

With a vehicle range of eight, there is far greater overlap between the segments generated by the mid-path algorithm and the nodes added by the ANDP. The ANDP sites in Fig. 6(b) are the same as the first 17 added in Fig. 6(a). The algorithms simply stopped after adding 17 instead of 22 nodes. The minimum separation distance achieved by the ANDP maximin is 2.5, found on those arcs of length five to which a node was added. The maximum separation achieved by the ANDP minimax is 4.5, found on the arc of length 9 to which one node had been added.

To summarize, the ANDP maximin is motivated by the idea of not wasting candidate sites by putting any too close together, while the minimax aims to avoid long arcs with no candidate sites, and the mid-path segment algorithms finds locations where one facility can refuel paths that would otherwise require a combination of two vertices. Depending on the range of the vehicle, the mid-path method can find locations similar to, or very different from, the ANDP methods. The ANDP maximin and minimax both favor longer arcs, but differ on how quickly they double back to add multiple nodes to the longest arcs.

6 Flow-refueling location model results

As explained earlier, we ran a total of 300 solutions (12×25) of the FRLM, which are displayed in the tradeoff curves of Fig. 7. Several general trends are evident in the graph across all vehicle ranges, p, and methods of adding candidate sites. First, the flow volume that can be refueled with a given p increases with the range of the vehicle. As the range increases, fewer refueling stops are needed on the round-trip journey.

[1] The overall maximin is 2.0, found on the original arcs of length two.

Second, for a given range, the refueled flow volume increases with the number of facilities built. The curves, however, can level off at a value less than 100% if there are no combinations of candidate sites and facilities that can refuel any of the remaining paths. Leveling off at a value below 100% is more likely if the range of the vehicle is small relative to the lengths of the arcs, and the number of candidate sites added to the arcs is few.

Third, the refuelable flow volume tends to increase at a decreasing rate, but the curve is not strictly convex. This is largely due to the need for a combination of facilities to refuel many paths. The curve may flatten out briefly if the next facility cannot refuel the new paths by itself, and then steepen again if the following facility completes one or more combinations that can refuel several paths. These first three patterns are expected conclusions consistent with those for the FRLM with vertex location only (Kuby and Lim 2005), but here we have confirmed that they continue to hold when candidate sites are added to arcs.

Moving on to results comparing the methods of adding candidate sites, the fourth general finding is that the ANDP methods for adding candidate sites on the arcs eventually outperform the mid-path segment method and the vertices alone as the number of facilities rises. The performance gap over vertices alone is largest for short-range vehicles and large p. For range=4, the gap first appears at 10 facilities, begins to widen at 16 facilities, and by p=25 grows to 20% on the vertical axis (or 1.29 times more flow refueled). For range=6, the gap becomes noticeable at p=18 and grows to 12–13% (1.15 times more flow). For

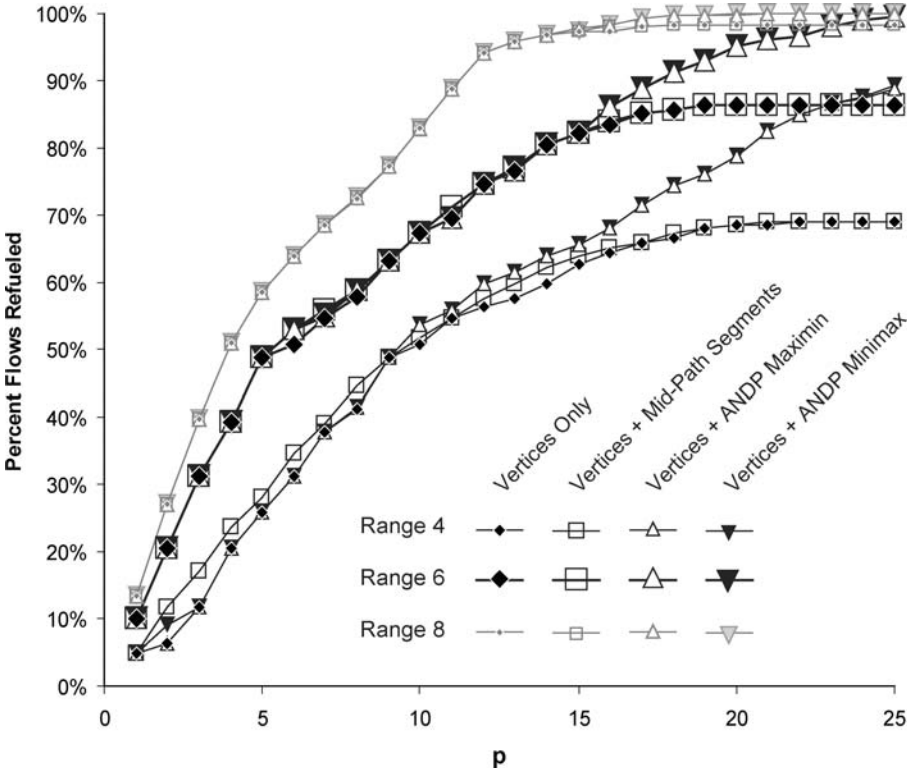


Fig. 7 Objective function value by vehicle range, number of facilities, and method for adding candidate sites on arcs

range=8, the gap appears around p=16, but never gets larger than 2%. Adding candidate sites by the ANDP method is thus more important for smaller ranges and larger number of facilities.

Figure 8 shows a typical solution in which the ANDP sites help achieve a better solution to the FRLM. This solution for a vehicle range of four locates 19 facilities at 15 vertices and four ANDP maximin sites. The ANDP sites on arcs 4–8 and 8–10, for example, can refuel trips amongst nodes 4, 8, and 10—but only in combination with facilities at the three nodes themselves. The mid-path segment algorithm, on the other hand, did not generate candidate sites on these arcs because they are too long for a single facility to refuel a vehicle with a range of only four.

Fifth, the FRLM tends to cluster facilities in the more populated and denser parts of the network, but not pack them any closer than necessary for the range of the vehicle. In Fig. 8, facilities are concentrated on the left side of the network, with no facilities located less than three or more than four distance units from another. Clustering them together enables them to work in combinations to refuel more paths. In addition, because shorter paths tend to have

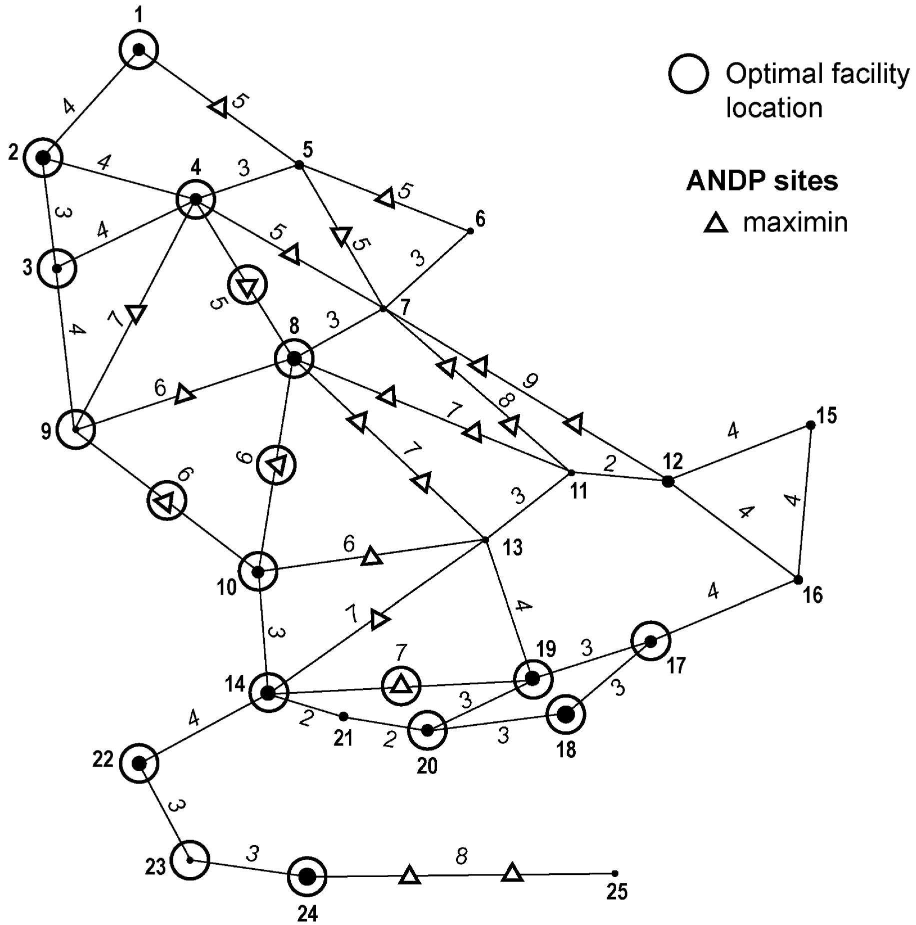


Fig. 8 Optimal solution for p=19 facilities, vehicle range=4, and 22 ANDP maximin candidate sites. Four facilities are located along arcs and 15 are located at vertices

heavier flows, clustering them together ensures that few of the short but heavy flows fall through the cracks. This strategy of grouping facilities near each other so they can work together in combinations is not often found in the location literature, and stands in marked contrast to the location strategies typical of median or covering models that spread facilities across a network.

Sixth, the mid-path segment candidate sites can sometimes outperform the ANDP candidate sites—but only for short-range vehicles and a small number of facilities. For a range of four, the FRLM solutions using 25 vertices plus the 22 mid-path segments outperform both ANDP methods and the vertices only from $p=2$ to $p=8$ by 3–53%. A good example of where the mid-path method excels is the $p=3$ solution for a range of four (Fig. 9). Two of the three facilities are on mid-path segments. These segments were originally generated because they can refuel the 14–22 and 18–20 round trips by themselves, but by spacing the facilities 3.5–4 distance units apart, the FRLM can use these stations in combinations to facilitate all possible trips among nodes 14, 18, 20, 21, and 22. Trips among these five nodes add up to 17.2% of all trip volumes—46% better than the ANDP solutions. The solutions for $p=4$ to $p=8$ build on this cluster by adding facilities to neighboring segments.

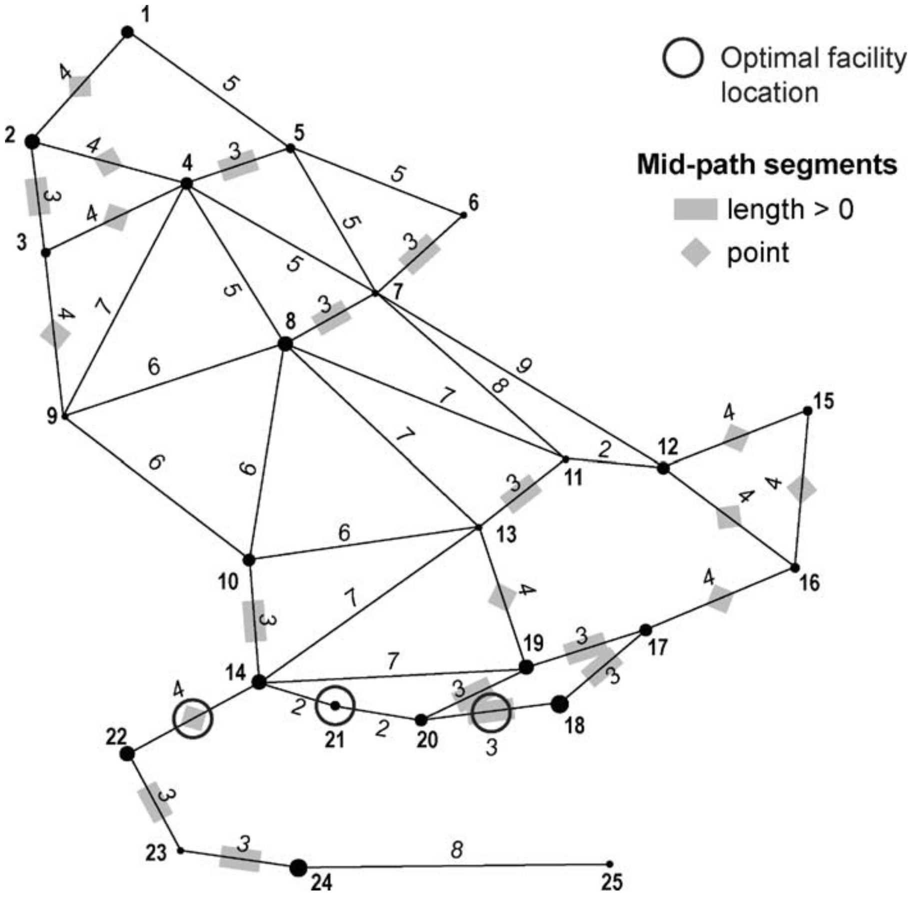


Fig. 9 Optimal solution for $p=3$ facilities, vehicle range=4, and 22 mid-path segment candidate sites

Generally speaking, the advantage of the mid-path segment method over the ANDP methods for shorter-range vehicles and smaller numbers of facilities appears to result from the placement of candidate segments on shorter arcs. With 17–22 added nodes, the ANDP method adds few nodes to shorter arcs. As p increases, it becomes less important to build facilities in the middle of shorter arcs, because the model can build enough facilities to place them at vertices on both ends of many shorter arcs. As p increases, the ANDP methods surge ahead by enabling the multi-arc trips that use some of the longer arcs.

Seventh, comparing the vertices plus maximin ANDP with the vertices plus minimax ANDP, the minimax comes out ahead by a tiny margin. Over 75 solutions (3 ranges and 25 p values), the minimax refuels 0.05% more flow volume than the maximin. The minimax sites outperform the maximin sites ten times, whereas the maximin does better on two solutions. As the range increases, the minimax's advantage decreases from 0.13% (range=4) to 0.02% (range=6) and then disappears altogether (−.01% for range=8). Although the differences are slight, it makes sense that the minimax would do slightly better for shorter-range vehicles by placing more candidate sites on shorter arcs, while the maximin would do slightly better for longer-range vehicles by more frequently placing multiple sites on longer arcs.

Vertex locations generally have an advantage over locations on arcs because a vertex can serve paths using any arc connected to the vertex in any direction, whereas a location on an arc can serve only those paths using that arc. In keeping with this principle, 14% of optimal facility locations are located on arcs, and 86% at vertices (Table 2). Figure 10 breaks down the percentage of facilities on arcs for each p value by (a) candidate site method (across all three ranges) and by (b) vehicle range (across all candidate site methods). Generally speaking, the percentage of facilities on arcs increases with p and decreases with the range. The percentage increases with p because arc locations are somewhat of a last resort for serving flows, and therefore tend not to be chosen until the good vertex locations are used up. The percentage decreases with vehicle range because longer-range vehicle have less need to refuel in the middle of arcs. Both of these trends are somewhat obscured, however, by the influence of the scenarios using mid-path segments and a range of four, in which the percentage of facilities on arcs was quite high (50–88%) for low p (2–8 facilities).

Table 2 Frequency of optimally locating facilities on arcs

Range	Method for adding sites on arcs	Percent of optimal locations on arcs
4	Vertices+mid-path segments	46
	Vertices+ANDP maximin	10
	Vertices+ANDP minimax	13
	All methods for range of 4	20
6	Vertices+mid-path segments	11
	Vertices+ANDP maximin	8
	Vertices+ANDP minimax	13
	All methods for range of 6	10
8	Vertices+mid-path segments	7
	Vertices+ANDP maximin	6
	Vertices+ANDP minimax	5
	All methods for range of 8	6
	All methods for all ranges	14

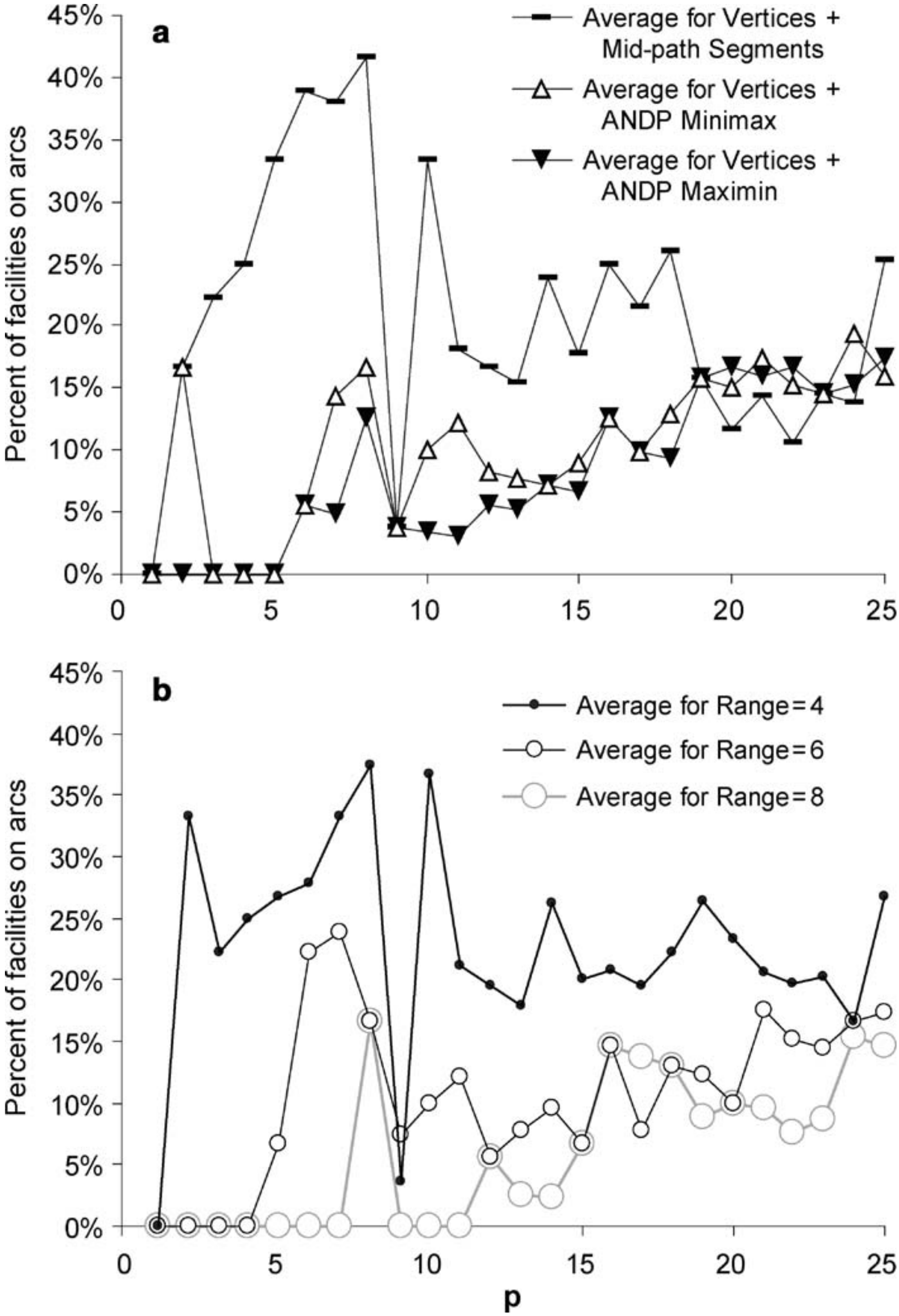


Fig. 10 Percentage of optimal facilities located on arcs, by (**a**) method of adding candidate sites (averaged over all ranges) and (**b**) vehicle range (averaged over all methods)

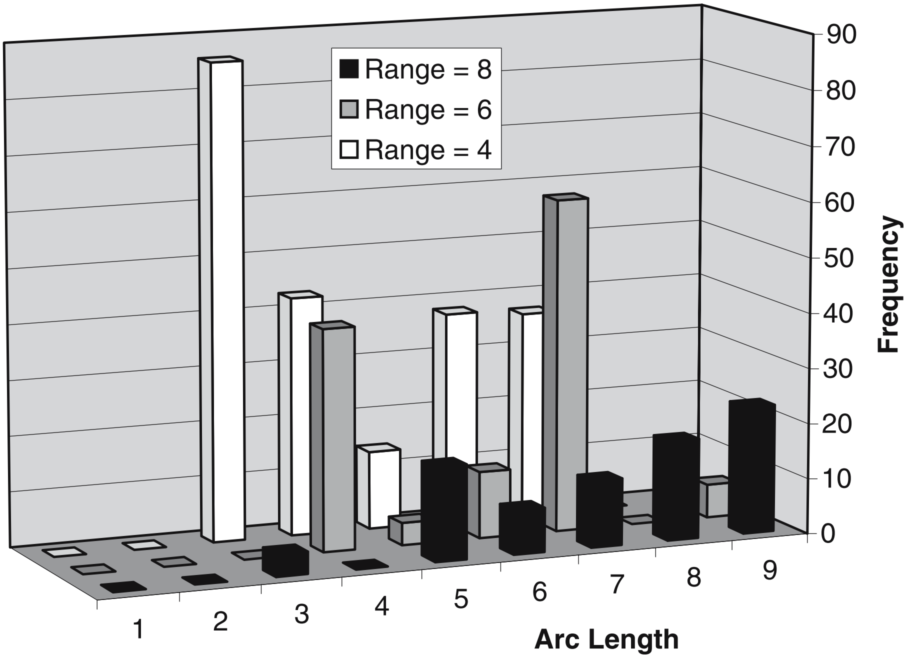


Fig. 11 Number of optimal facilities located on arcs, by arc length. These frequencies represent the total for all three methods of locating candidate sites along arcs

As Fig. 10 shows, candidate sites on arcs make up a substantial percentage of the optimal locations for all ranges, all methods of adding candidate sites, and low and high numbers of facilities. It does not appear that analysts can afford to ignore locations on arcs for any subset of problems.

Because of the inherent advantage of a vertex location, it stands to reason that the farther a vehicle can go between refuelings, the less it will need to stop to refuel between vertices. As a result, it is not surprising that more facilities tend to be located on longer arcs as vehicle range increases (Fig. 11).[2] With a range of eight, vehicles will need only one stop amid even the longest arcs, whereas with a range of four it may take two facilities placed on a long arc. In addition, in our case study the flow volumes are a function of a gravity model process in which shorter-distance flows will tend to have higher volumes than long-distance flows. Thus, with short-range vehicles, optimal locations tend to be at vertices or on shorter arcs that refuel short paths over short arcs that contribute more to the objective function *and* are cheaper to cover. But as the range increases, these same short trips are easily refueled by facilities at vertices, and the non-vertex locations shift towards longer arcs.

As Figs. 7, 8, and 9 showed, adding candidate sites to the arcs of the network is crucial for getting good solutions to the FRLM, but no one method of adding candidate sites consistently outperforms the others for the same number of added sites. While the ANDP sites consistently outperformed the mid-path segments *for large p and longer range*, the

[2] Comparison between arcs of different lengths is somewhat problematic because the network contains unequal numbers of each arc length, and some arcs are more centrally located than others. Nevertheless, the network is constant across all runs, so the comparisons between ranges 4, 6, and 8 are still valid.

segments did better for small p and a range of four. The reason is simple—the segment method placed sites on short arcs, while the ANDP would have had to add more sites than allotted to reach those same smaller arcs. To confirm this, we increased the number of ANDP sites from 22 to 50 to 75. Table 3 shows that the vertices plus 75 ANDP sites were never outperformed for any number of facilities. The 22 added segments obtained equally high objective function values for $p=1$ to 9, and the 50 added ANDP sites did the same for $p=1$ and 9–25.

The brute force method of adding more and more ANDP sites, however, is not a viable strategy. Even with this small sample problem, the computational burden begins to explode. As Table 4 shows, increasing the ANDP sites from 50 to 75 increases: the number of binary variables x_k from 75 to 100; the number of combination variables v_h from 1,468 to 11,031; the number of rows from 9,720 to 106,575; and the CPU time for solving $p=1$ to 25 from 4 minutes to 21 hours.

7 Conclusions

The flow-refueling location model was designed to locate p refueling stations so as to refuel the maximum volume of traffic flows traveling on their shortest paths from origins to

Table 3 Percent of trips refueled, by candidate site method and number of added sites. Best results achieved for each p value are highlighted. Range=4

Number Of Facilities	Vertices Only	Vertices + 22 Mid-Path Segments	Vertices + 22 ANDP Maximin	Vertices + 50 ANDP Maximin	Vertices + 75 ANDP Maximin
1	4.9	4.9	4.9	4.9	4.9
2	6.3	11.8	6.3	9.1	11.8
3	11.8	17.2	11.8	12.0	17.2
4	20.4	23.7	20.4	20.4	23.7
5	25.9	28.0	25.9	25.9	28.0
6	31.3	34.7	31.3	31.3	34.7
7	37.8	39.0	37.8	37.8	39.0
8	41.3	44.7	41.4	41.4	44.7
9	48.8	48.8	48.8	48.8	48.8
10	50.7	51.7	53.6	53.6	53.6
11	54.6	54.6	56.0	57.7	57.7
12	56.3	57.5	59.8	60.6	60.6
13	57.5	59.9	61.5	63.9	63.9
14	59.9	62.1	63.9	66.8	66.8
15	62.6	63.8	65.6	69.1	69.1
16	64.3	65.2	68.1	70.8	70.8
17	65.9	65.9	71.6	72.3	72.3
18	66.5	67.3	74.4	75.7	75.7
19	67.9	67.9	76.1	78.5	78.5
20	68.6	68.6	78.7	80.2	80.2
21	68.6	69.0	82.4	82.4	82.4
22	69.0	69.1	84.9	84.9	84.9
23	69.1	69.1	86.6	86.6	86.6
24	69.1	69.1	87.4	87.6	87.6
25	69.1	69.1	88.9	89.2	89.2

Note: Best results achieved for each p value are highlighted. Range = 4.

Table 4 Computational performance for different numbers of ANDP maximin sites

Number of					CPU time for	
ANDP maximin sites	Binary x_k variables	Continuous v_h combination variables	Rows	Columns	Generating ANDP sites and combinations	Solving the FRLM for p=1 to 25
0	25	73	219	102	01 s	10 s
22	47	280	1,588	341	10 s	12 s
50	75	1,468	9,720	1,801	14 s	4 min:09 s
75	100	11,031	106,575	11,404	7 min:13 s	21 h:32 min:40 s

Note: Range=4.

destinations. The FRLM extends the flow-capturing/intercepting problem by incorporating the limited driving range of vehicles, which leads in turn to the need for combinations of facilities to refuel longer paths and to the vertices not being a finite-dominating set. This paper introduced and compared three methods for generating additional discrete candidate sites along the arcs of a network for obtaining better solutions to the FRLM with the same number of facilities.

Several findings of this paper—such as clustering of optimal locations and convexity (but not strict convexity) of the objective function with respect to p—were consistent with earlier results that did not add candidate sites to arcs. The key new finding of this paper is that, despite the inherent advantage of vertex locations, adding candidate sites along arcs substantially improves the coverage of the network compared with considering only locations at vertices. The improvement was often slight for small numbers of facilities, but as the number of facilities increases, the objective function obtained using additional ANDP candidate sites gradually leaves the objective obtained using network vertices only further and further behind. This decided advantage of adding ANDP sites is highly pronounced for short-range vehicles but diminishes with longer-range vehicles.

The ANDP methods add sites first to the longest arcs and then work their way to the shorter arcs, with the minimax ANDP reaching the shorter arcs a little sooner than the maximin. A similar generalization is not possible for the mid-path segment method, because the locations depend on the vehicle range assumed. Assuming a short range, segments tend to be on short arcs, in marked contrast to the ANDP sites. Assuming longer range, however, there is considerable overlap with the ANDP sites.

The tendencies of the ANDP and segment methods to place candidate sites on different length arcs greatly affected their ability to improve FRLM results. Although the ANDP methods achieved pronounced, consistent improvements of the objective function over the vertices for large p, the mid-path segments cannot be discounted. Segments performed better than ANDP sites for short-range vehicles and a small number of facilities when the same number of sites was added. Given a greater number of candidate sites to work with, the ANDP methods eventually begin to locate sites on the shorter arcs and obtain equal or better FRLM results than the segment method for all p—but the computational cost grows exorbitantly high.

While arcs shorter than half the vehicle range can be safely ignored, our results indicate that it is important to add candidate sites to all other arcs. Neither medium arcs (from 50–100% of the range) nor long arcs (greater than the range) can be ruled out as optimal locations. As a result, neither the mid-path segment method nor the ANDP method is superior to the other in all circumstances. Because the mid-path segments are selected as specific pieces of specific arcs that can refuel an entire path with a single non-vertex facility, they sometimes do better than the ANDP sites for few facilities and short-range vehicles. However, because the segment algorithm does not take into account how they will work in combinations, nor attempt to spread them around, they do not improve the objective function very much for larger numbers of facilities. In contrast, the ANDP sites are not chosen with any particular paths in mind, but by spreading them around the network away from vertices and each other, they are able to coordinate with vertices and each other in combinations in unforeseen ways.

More research is necessary to explore other approaches that efficiently combine the best features of the ANDP and mid-path segment methods. The segment method could be extended to identify segments that would work in combination with vertices to refuel longer paths with fewer facilities than otherwise needed. Alternatively, one could apply the mid-path segment method first, and then use the ANDP method to disperse candidate sites away from vertices, segment midpoints, and each other.

If government agencies or energy companies begin using the FRLM to locate refueling stations for hydrogen or other alt-fuel vehicles on networks at the state or national scale, it will be important for them to generate additional candidate sites. This will be especially true in areas where the length of arcs is long relative to the range of the vehicle, as in the US Great Plains and Rocky Mountains. Short arcs, however, cannot be ignored when adding candidate sites. Based on the solutions with vehicle range of eight, in which the vertices-only solution was nearly as good as that with vertices plus 17 ANDP sites, additional candidate sites on arcs most likely will be irrelevant on urban networks. In modeling urban networks, however, researchers will need to pay close attention to how to model freeway intersections. Drivers cannot usually exit to a filling station at an intersection of two limited-access highways. One way to model these intersections might be to establish candidate sites at the freeway exits closest to the actual interchange. Another approach would be to consider the closest exits to be part of the freeway interchange itself, and use signage to direct drivers to a single exit containing the alt-fuel refueling station.

Other promising avenues beckon for future research on locating refueling stations for range-limited vehicles. Upchurch et al. (2007) extended the FRLM to capacitated facilities. We are also testing heuristic solution algorithms, which will be needed as researchers attempt to apply the FRLM to large networks and add large numbers of candidate sites to arcs. Kim and Kuby (2006) have done preliminary research on modeling driver’s potential detours from their shortest paths to reach a refueling station.

Although the added-node dispersion problem is not designed to generate a finite dominating set for other location problems, it has now significantly improved the results of two different discrete location problems for which finite-dominating sets have not been determined. Future work should test the ANDP on other facility location problems lacking finite dominating sets. The ANDP is Order(mp) where m is the number of original arcs and p is the number of added nodes, so it solves quickly and is simple to program. It would be interesting to compare the performance of the ANDP with the globally optimal solution for a problem in which the finite-dominating set is known.

Acknowledgments This research was supported by the Decision, Risk, and Management Science program of the National Science Foundation, Proposal #0214630. We would like to thank Dash Optimization, Ltd. for the use of Xpress-MP under the Academic Partnership Program. We also thank Max Wyman for his inspiration to investigate location of refueling stations for hydrogen fuel-cell locomotives, and Barbara Trapido-Lurie for her graphical expertise. We also thank the anonymous reviewers, whose comments led to substantial improvements.

References

Bapna R, Thakur LS, Nair SK (2002) Infrastructure development for conversion to environmentally friendly fuel. Eur J Oper Res 142:480–496

Berman O, Simchi-Levi D (1988) A heuristic algorithm for the traveling salesman location problem on networks. Oper Res 36:478–484

Berman O, Larson RC, Fouska N (1992) Optimal location of discretionary service facilities. Transp Sci 26:201–211

California Environmental Protection Agency (2005) California hydrogen blueprint plan, vol 1. May 2005

Church RL, Meadows ME (1979) Location modeling using maximum service distance criteria. Geogr Anal 11:358–373

Church RL, ReVelle CS (1974) The maximal covering location problem. Pap Reg Sci Assoc 32:101–118

Daskin M, (1995) Network and discrete location: models, algorithms, and applications. John Wiley & Sons, New York

Erkut E (1990) The discrete p-dispersion problem. Eur J Oper Res 46:48–60

Goodchild MF, Noronha VT (1987) Location–allocation and impulsive shopping: the case of gasoline retailing. In: Ghosh A, Rushton G (eds) Spatial analysis and location–allocation models. Van Nostrand Reinhold, New York

Hodgson MJ (1990) A flow capturing location–allocation model. Geogr Anal 22:270–279

Hooker JN, Garfinkel RS, Chen CK (1991) Finite dominating sets for network location problems. Oper Res 39:100–118

Kim J-G, Kuby M (2006) Locating refueling stations for alternative fuel vehicles on detouring paths. Paper presented at the Association of American Geographers Annual Meeting, Chicago, Illinois, 8 March 2006

Kuby M (1987) Programming models for facility dispersion: the p-dispersion and maxisum dispersion problems. Geogr Anal 19:315–329

Kuby M, Lim S (2005) The flow-refueling location problem for alternative-fuel vehicles. Socio-Econ Plann Sci 39:125–145

Kuby M, Lim S, Wang K (2004) A model for optimal location of hydrogen refueling stations: an Arizona case study. Hydrogen: a clean energy choice (Proceedings of the National Hydrogen Association's 15th annual U.S. hydrogen conference and hydrogen expo USA)

Kuby MJ, Lim S, Upchurch CJ (2005) Dispersion of nodes added to a network. Geogr Anal 37:384–409

Melaina MW (2003) Initiating hydrogen infrastructures: preliminary analysis of sufficient number of initial hydrogen stations in the US. Int J Hydrogen Energy 28:743–755

Melendez M, Milbrandt A (2005) Analysis of the hydrogen infrastructure needed to enable commercial introduction of hydrogen-fueled vehicles. The Proceedings of the National Hydrogen Association, 29 March 2005

Nicholas M, Handy S, Sperling D (2004) Using geographic information systems to evaluate siting and networks of hydrogen stations. Transp Res Rec 1880:126–134

Toregas C, ReVelle CS (1973) Binary logic solutions to a class of location problems. Geogr Anal 5:145–155

Upchurch C, Kuby M, Lim S (2007) A capacitated model for location of alternative-fuel stations. Geogr Anal (in press)