Analysis of Errors Due to Demand Data Aggregation in the Set Covering and Maximal Covering Location Problems

In this paper, we extend the concepts of demand data aggregation error to location problems involving coverage. These errors, which arise from losses in locational information, may lead to suboptimal location patterns. They are potentially more significant in covering problems than in p-median problems because the distance metric is binary in covering problems. We examine the Hillsman and Rhoda (1978) Source A, B, and C errors, identify their coverage counterparts, and relate them to the cost and optimality errors that may result. Three rules are then presented which, when applied during data aggregation, will reduce these errors. The third rule will, in fact, eliminate all loss of locational information, but may also limit the amount of aggregation possible. Results of computational tests on a large-scale problem are presented to demonstrate the performance of rule 3.

INTRODUCTION

The set covering location problem (Hakimi 1964; Toregas and ReVelle 1972) and the maximal covering location problem (Church and ReVelle 1974) have received considerable attention in the facility location literature due to their widespread applicability. As is reported in Current and Storbeck (1988) these models have been applied to the location of day care facilities, fire stations, bus stops, emergency medical services, social service centers, airports, and archaeological settlement analysis, among others.

The basic underlying assumption of these models is that there exists some covering distance, S, such that demand is satisfied, or covered, if it is within S distance of a facility, and that demand is not satisfied or covered if it is not within S distance of a facility. The objective of the set covering location problem (SCLP) is to locate the minimum number of facilities subject to the restriction that all demand must be covered. Given that the minimum number of facilities necessary to cover the entire demand may be infeasible or impractical for budgetary reasons, the objective of the maximal covering location problem (MCLP) is to locate a fixed number of facilities, p, in such a way as to cover as much of the total demand as possible.

John R. Current is associate professor and David A. Schilling is professor of management science, Ohio State University.
The formulations for these problems appear in Toregas and ReVelle (1972), Church and ReVelle (1974), and Current and Storbeck (1988) and consequently will not be given here. Another underlying assumption of these models, however, is that demand may be represented as existing at a finite number of nodes and that there exist a finite number of potential facility sites which may also be represented by nodes.

Unfortunately, the SCLP (Karp 1972) and the MCLP [through its relationship to the p-median problem, Church and ReVelle (1976) and Carey and Johnson (1979)] belong to the class of problems known as NP-hard; therefore, it is unlikely that there exist efficient polynomially bound algorithms to obtain optimal solutions for them. Due to the computational complexity of the SCLP and the MCLP several heuristic solution approaches have been developed (e.g., Vasko and Wolf 1988; Church and ReVelle 1974). However, even with the utilization of specialized heuristics, large problem instances of the SCLP and MCLP may be difficult to solve in reasonable time, especially on personal computers. As a consequence, analysts frequently aggregate the demand data to reduce the number of demand nodes and thereby reduce the problem size and computational requirements.

As has been demonstrated for the p-median location problem (Hillsman and Rhoda 1978; Goodchild 1979; Casillas 1987, Current and Schilling 1987), aggregation of demand data results in the loss of locational information which, in turn, may produce suboptimal locational patterns. The objective of the p-median location problem (PMP) is to locate p facilities in such a way as to minimize the total travel distance that demand must traverse to reach its nearest facility. Consequently, the PMP measures the actual (or approximated) distances that demand must travel.

Actual distances in the SCLP and MCLP are not directly considered, but rather are converted to a binary metric (Church and ReVelle 1976). More specifically, distances in these two problems may be viewed as equalling zero if the actual distance is less than or equal to S, and infinity if not. Consequently, errors in distance measurement caused by demand data aggregation may be much more significant in the SCLP and MCLP than they are in the PMP. A small error in the measurement of the actual distance may lead to the difference between zero and infinity in the converted distance metric. As cited above, several articles have analyzed the effects of demand data aggregation on the PMP. Prior to this paper, however, none have analyzed the effects of such errors on the SCLP and only Daskin et al. (1989) have analyzed these effects on the MCLP.

Hillsman and Rhoda (1978) described three sources of errors resulting from demand data aggregation in the p-median problem (Source A, B, and C errors). In this paper, we define and analyze the effects of these errors on the SCLP and the MCLP. In addition, we demonstrate how they may be avoided in planning situations where the analyst performs the aggregation.

The rest of this paper is organized as follows. In the next section, we define the errors for covering problems, discuss their potential impact, and show that one of these error classifications is not relevant in the SCLP and MCLP. In the third section, we describe methods to eliminate the other two sources of errors. In the fourth section we present test problem results. In the final section, we discuss conclusions and present a summary of our findings.

DESCRIPTION OF SOURCE A, B, AND C ERRORS IN LOCATION COVERING MODELS

In describing these error sources, we will employ the notation used in Current and Schilling (1987). Specifically, the nodes representing unaggregated demand (that is, their real locations) are referred to as basic spatial units (BSUs) and the
nodes representing the aggregated demand are referred to as aggregated spatial units (ASUs). Basic spatial units will be indexed by the subscript $k$ (BSU$_k$), aggregated spatial units will be indexed by the subscript $i$ (ASU$_i$), and potential facility sites (PFSs) will be indexed by the subscript $j$ (PFS$_j$).

Casillas (1987) and Current and Schilling (1987) describe two types of errors that result from three error sources in the p-median problem. These are optimality errors and cost errors. A cost error in the PMP is the difference between the measured cost (the objective function value) of a solution and the true cost (the actual weighted travel distance) for that solution. The SCLP and the MCLP do not consider the actual distance between demand and a facility except to determine whether or not it is less than or equal to the covering distance. Consequently, we will now refer to this type of error as the covering error rather than the cost error. Specifically, the covering error is the difference between the solution's stated covering and the true covering of the disaggregated data for that solution.

The optimality error in the PMP is the difference between the true cost (that is, having corrected for cost errors) of an aggregated solution and the cost of the optimal solution to the unaggregated version of the problem. The optimality error, therefore, measures the effect of locational changes caused by aggregation. This definition will serve covering problems, as well. These definitions are also consistent with those used in Daskin et al. (1989) for the MCLP.

Error source A for the p-median problem is demonstrated in Figure 1. Let $u_k$ and $u_{k+1}$ represent the demands at BSUs $k$ and $k + 1$, respectively (assume no demand at $i$); $a_i$ equals $u_k + u_{k+1}$; and $d_{mj}$ equals the distance between node $m$ and $j$, $m$ equals $k$, $k + 1$, $i$. In this example demands at BSUs $k$ and $k + 1$ have been aggregated at ASU$_i$.

If ASU$_i$ is assigned to PFS$_j$ in the PMP, then $a_id_{ij}$ would represent the aggregated travel distance for this assignment. The true (or unaggregated) distance for this allocation, however, is $u_kd_{k+1} + u_{k+1}d_{k+1,i}$. The difference between these two distance values represents the Source A cost error in the PMP. As is described in Current and Schilling (1987) this error source may lead to a positive cost error (overestimate the true weighted travel distance) or a negative one (underestimate the true travel distance) and may cause optimality error as well. Current and Schilling (1987) present a measurement scheme which eliminates this error source in the PMP.

There are two cases of Source A errors in the SCLP and MCLP. In the first, a BSU is considered as covered when in fact it is not; and in the second a BSU is covered but is not counted as such. Consider Figure 2, where a facility is located at node $j$ which covers all demand nodes within the line encircling node $j$. BSU$_k$ is not covered by a facility at node $j$. However, the demand at BSU$_k$ has been
aggregated at \( ASU_j \) which is covered by a facility at node \( j \). Therefore, if a facility is located at node \( j \), both the SCLP and MCLP will consider the demand at \( BSU_k \) to be covered by the facility at node \( j \), even though this is not true. We refer to this type of Source A error as Case 1.

The second type (Case 2) of Source A errors is also demonstrated in Figure 1 where \( BSU_{k+1} \) which is within the covering distance, \( S \), of node \( j \) is aggregated to \( ASU_{j+1} \) which is not within \( S \) distance of node \( j \). Consequently, if a facility is located at node \( j \), neither the SCLP nor the MCLP will count \( BSU_{k+1} \) as being covered by the facility at node \( j \), even though it is.

Source A, Case 1 errors in the SCLP and MCLP will always result in non-negative covering errors. That is, if such an error exists, it will be because the solution reports that it covers more demand than it actually does cover. This error may be zero, however, even if Case 1 aggregation occurs. For example, in Figure 2, another facility may be sited at node \( j + 1 \) which does cover \( BSU_k \).

Source A, Case 2 errors will always result in nonpositive covering errors. That is, if such an error exists, it will be because the solution reports that it covers less demand than it actually does. Again, this error may be zero even if Case 2 aggregation occurs. For example, in Figure 2, another facility may be sited at node \( j + 2 \), which covers \( ASU_{j+1} \); therefore, \( BSU_{k+1} \) will be counted as covered. Combined Source A errors may be positive or negative. Consequently, the overall direction of the Source A covering error cannot be determined without examining the solution vis-à-vis the underlying unaggregated data.

The optimality errors for the SCLP and MCLP must be discussed separately because the two problems have different objectives. In the SCLP, Source A, Case 1 aggregation errors may result in infeasible solutions to the unaggregated problem because some BSUs may not be covered even when all of the ASUs are covered. Source A, Case 2 aggregation errors may result in a solution to the SCLP which requires more facilities than necessary to cover all the BSUs. This can only happen, however, if a BSU is aggregated at an ASU which has no "natural" demand located at it (that is, it is not the site of a BSU). For example, in Figure 2, if \( ASU_{j+1} \) has no natural demand located at it, it will need to be covered by a facility to cover \( BSU_{k+1} \)'s demand even though that demand is really covered by a facility at node \( j \). If, however, \( ASU_{j+1} \) has some demand of its own then it will need to be covered whether \( BSU_{k+1} \)'s demand is aggregated to it or not.

Optimality errors in the MCLP resulting from Source A errors may also be significant. For example, a potential facility site, \( j \), which has a high percentage of its total potential covering resulting from Case 1 errors may be selected over
another potential site, \( j + 1 \), which has lost a high percentage of its true covering resulting from Case 2 errors. The selection of site \( j \) rather than \( j + 1 \) may yield a large optimality error if node \( j + 1 \) covers more unaggregated demand than does node \( j \).

Source B errors in the \( p \)-median problem occur when a facility is located at an ASU and may be viewed as special cases of Source A errors. Consider Figure 3 where demands at \( BSU_k \) and \( BSU_{k+1} \) have been aggregated to ASU\(_j\). In addition, a facility has been sited at ASU\(_j\). In such a solution to the PMP, the weighted travel distance for this allocation would be zero. The true weighted travel distance, however, would be \( u_k d_{k} + u_{k+1} d_{k+1} \). The difference between these two values is the Source B cost error. As described in Current and Schilling (1989) this cost error will always be nonpositive (that is, underestimate the true cost) and may yield optimality error. A measurement scheme for eliminating Source B error in the PMP is presented in Current and Schilling (1987).

Source B errors are similar in the SCLP and MCLP. For example, in Figure 4, if a facility is located at ASU\(_j\), the demand at BSU\(_k\) which has been aggregated to ASU\(_j\) will be counted as covered when, in fact, it is not. Consequently, the covering errors resulting from Source B errors will be positive or zero. These errors can be eliminated during aggregation by only aggregating the demand at a BSU to an ASU if the distance between them is less than or equal to the covering distance, \( S \) (aggregation rule 2, discussed later). The potential optimality errors resulting from Source B errors for the SCLP and MCLP are similar to those for Source A Case 1 errors and will not be discussed again, except to state that they can be eliminated from the SCLP if the above aggregation rule is followed and if only BSUs are considered as potential ASUs.
Source C errors in the p-median problem are demonstrated by Figure 5. In the PMP, demand at BSU_k will be assigned to the facility at node j even though it is closer to the facility at node j + 1 because the demand has been aggregated at ASU_i. As described in Current and Schilling (1987), Source C errors may result in both cost and optimality errors with the cost errors always being non-negative.

Source C errors in the PMP result from assigning a BSU to an open facility which is not the nearest open facility. In the SCLP and MCLP demand can only be assigned to an open facility which covers it. The actual distance separating demand from a facility is not measured, as long as it is less than S. Therefore, no analogue to Source C errors in the PMP exists for the SCLP or the MCLP.

Daskin et al. (1989) examined the error effects for three aggregation schemes for the MCLP. Their results indicate that the aggregation scheme employed may have significant influence on the resulting errors. Their aggregation schemes selected ASUs from the set of BSUs based on (1) demand at the BSUs; (2) distance between the ASUs; and (3) demand weighted distance between the ASUs. The overall average covering errors for the various schemes ranged from 4.69 percent to 10.79 percent with the maximum error being 37 percent. The overall average optimality errors for the three schemes ranged from 3.78 percent to 15.01 percent with the maximum error being 28.90 percent. However, two of the schemes, discussed later, performed considerably better than did the third.

None of the aggregation schemes analyzed by Daskin et al. (1989) were directly designed to eliminate the error sources discussed in this section. In the next section, we propose three aggregation rules (two of which were introduced above) which are specifically designed to reduce the effects of the error sources.

METHODS TO ELIMINATE SOURCE A AND B ERRORS

As was demonstrated in the previous section, two straightforward aggregation rules will eliminate some of the errors associated with the aggregation of demand...
data in the SCLP. These rules may be stated as follows:

1. Only consider BSUs as potential ASUs.
2. Do not aggregate the demand at \( BSU_k \) to \( ASU_i \) if \( d_{k,i} > S \), where \( d_{k,i} \) is the distance between \( BSU_k \) and \( ASU_i \) and \( S \) is the covering distance.

The implementation of these two rules will have the following effects on aggregation errors associated with the SCLP:

a) Source A, Case 2 errors will not lead to covering or optimality errors. Source A, Case 1 errors, however, may still lead to non-negative covering errors and an infeasible solution for the underlying unaggregated problem.

b) Source B errors will be eliminated.

The implementation of rule 2 in MCLPs will eliminate the effects of Source B errors. The Source A errors may result in positive or negative covering errors and optimality errors.

We now introduce a third aggregation rule which will eliminate the effects of all aggregation error types in both the SCLP and the MCLP. This rule may be stated as follows:

3. Only aggregate demand at a BSU to an ASU if the set of potential facility sites which cover each of them are identical.

Implementation of rule 3 will guarantee that if any facility is opened which covers an ASU it will also cover all of the BSUs which are aggregated at it. Obviously, if rule 3 is implemented, the conditions of rule 2 will also be satisfied. Unfortunately, utilization of this rule may greatly reduce the amount of aggregation possible. Nevertheless, rule 3 produces an aggregation pattern with no loss of locational information.

The methodology to implement rule 3 is as follows:

Step 0: Place all BSUs on the Active List.

Step 1: Select the first BSU on the Active List as the “template” BSU and assign it to a newly created ASU cluster.

Step 2: Find all remaining BSUs on the Active List that are covered by the same set of PFSs as the template BSU.
   a) Remove them from the Active List and
   b) Assign them to the current ASU cluster.

Step 3: If the Active List is empty—STOP. Otherwise, go to Step 1.

The essence of rule 3 is to collapse into ASU clusters all BSUs which are covered by an identical set of PFSs. In this way, each BSU in a given cluster (ASU) has the same covering characteristics as the ASU and, therefore, the ASU can act as a perfect covering representation of the BSUs it contains. Thus, no locational information is lost. As noted, however, a price is paid for this perfection in the amount of aggregation that can be achieved. The number of ASUs created by applying rule 3 is a function of the number and location of the PFSs, and the maximum covering distance (\( S \)). In the worst case, however, the number of ASUs created can be no larger than the number of BSUs (i.e., no aggregation at all). This would only be the case if each BSU was covered by a unique set of PFSs.

In practice, one expects that the level of aggregation will be high (the number of ASUs is small) when \( S \) is either small or large. When \( S \) is larger than the greatest BSU-to-PFS distance, all BSUs are covered by all PFSs and can be aggregated to a single ASU and the resulting problem is trivial. On the other hand, when \( S \) is less than the smallest BSU-to-PFS distance, no BSUs are covered except those at PFSs. Then, the number of ASUs is equal to the number of PFSs (all BSUs not at PFSs
cannot be covered and, therefore, are discarded). As $S$ moves away from either of these limits, the number of ASUs required to represent all BSUs should increase.

Rule 3 is similar to the row reduction process (rule 1) presented in Toregas and ReVelle (1973) as a solution procedure for set covering location problems. In fact, rule 3 presented here is more restrictive (that is, can, at best, equal the reductions obtained and generally yields less reduction in problem size) than the earlier technique. Consequently, the method presented in Toregas and ReVelle (1973) is recommended over rule 3 for SCLPs. However, the row reductions technique is based on the assumption that all demand nodes must be covered. When this assumption is relaxed as it is in many covering problems, their method cannot be used. For example, the technique cannot be used for the MCLP. Rule 3 retains all of the locational information present in the original network and is, therefore, suitable for any covering model with single maximal service distances.

TEST PROBLEM RESULTS

To gain a better understanding of the level of aggregation achievable, we applied rule 3 to the 681-node location network representing Baltimore City, Maryland, presented in Schilling et al. (1980). Six sets of PFSs were used. In the first, thirty BSUs were visually selected to be PFSs so as to be uniformly distributed across the network. For the second set, seventy BSUs were visually selected to be uniformly distributed PFSs. The third and fourth sets of PFSs were the thirty and seventy BSUs, respectively, that were identified as medians from the original 681-node network. These two sets, MED–30 and MED–70, together with the first two (UNI–30 and UNI–70) were obtained from Current and Schilling (1987). The final two sets of PFSs were thirty and seventy randomly selected BSUs (RAND–30, RAND–70).

Rule 3 was applied for each PFS set and for thirteen values of $S$, ranging from one-half mile to fifteen miles. (The maximum internode distance was approximately thirteen miles.) The results are presented in Table 1. As expected, the greatest aggregation occurs at the smallest and largest covering distances. Even at worst, however, the number of demand nodes could be reduced by 19 percent. It is worth noting that the entire 681-node network was originally used in a fire station siting study. For the covering distances used there, 1.0 to 2.0 miles, reductions of 36 to 92 percent in the number of demand nodes would have been possible with no loss in locational information.

<table>
<thead>
<tr>
<th>Covering Distance, $S$</th>
<th>UNI–30</th>
<th>UNI–70</th>
<th>OPT–30</th>
<th>OPT–70</th>
<th>RAND–30</th>
<th>RAND–70</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>31</td>
<td>74</td>
<td>31</td>
<td>102</td>
<td>42</td>
<td>116</td>
</tr>
<tr>
<td>1.0</td>
<td>57</td>
<td>312</td>
<td>79</td>
<td>250</td>
<td>87</td>
<td>242</td>
</tr>
<tr>
<td>1.5</td>
<td>149</td>
<td>450</td>
<td>150</td>
<td>396</td>
<td>147</td>
<td>361</td>
</tr>
<tr>
<td>2.0</td>
<td>185</td>
<td>415</td>
<td>205</td>
<td>455</td>
<td>204</td>
<td>458</td>
</tr>
<tr>
<td>3.0</td>
<td>237</td>
<td>483</td>
<td>266</td>
<td>516</td>
<td>289</td>
<td>498</td>
</tr>
<tr>
<td>4.0</td>
<td>244</td>
<td>525</td>
<td>300</td>
<td>554</td>
<td>253</td>
<td>304</td>
</tr>
<tr>
<td>5.0</td>
<td>244</td>
<td>525</td>
<td>282</td>
<td>536</td>
<td>240</td>
<td>305</td>
</tr>
<tr>
<td>6.0</td>
<td>214</td>
<td>479</td>
<td>202</td>
<td>438</td>
<td>181</td>
<td>440</td>
</tr>
<tr>
<td>7.0</td>
<td>167</td>
<td>373</td>
<td>129</td>
<td>304</td>
<td>121</td>
<td>332</td>
</tr>
<tr>
<td>8.0</td>
<td>98</td>
<td>236</td>
<td>74</td>
<td>175</td>
<td>63</td>
<td>211</td>
</tr>
<tr>
<td>9.0</td>
<td>55</td>
<td>183</td>
<td>40</td>
<td>91</td>
<td>31</td>
<td>123</td>
</tr>
<tr>
<td>10.0</td>
<td>19</td>
<td>49</td>
<td>18</td>
<td>34</td>
<td>13</td>
<td>37</td>
</tr>
<tr>
<td>15.0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Significant computation time advantages can accrue through this aggregation process. As an example, comparisons were made for the time needed to solve aggregated and unaggregated maximal covering location problems for two of the PFS sets: UNI-US and UNI-70. Using a realistic coverage distance of 2.0 miles, it took ten seconds to identify the 185 ASUs for UNI-US and fifty-four seconds to find the 415 ASUs of UNI-70. This "aggregation cost" compares quite favorably with the solution times for the aggregated versus unaggregated problems. It required 25 seconds to solve the 30-PFS, 185-demand node aggregated problem as opposed to 220 seconds to solve the 30 x 681-node original problem. For the 70 x 415 aggregated problem, 143 seconds were needed to find the solution, while 314 seconds were needed for the unaggregated 70 x 681-node model. In each case, the MIP package, SCICONIC, was used to produce the optimal solutions.

As stated earlier, a disadvantage of rule 3 is that the underlying data may limit the amount of aggregation possible. In the aggregation schemes proposed by Daskin et al. (1989), the number of ASUs can be determined in advance. If the implementation of rule 3 does not yield a sufficiently small number of ASUs, we recommend that Scheme A or Scheme C proposed in Daskin et al. (1989) be applied to the ASUs generated by rule 3. In addition, the Daskin et al. (1989) scheme employed should be modified by the inclusion of rule 2, when feasible. The Daskin et al. (1989) schemes may be summarized as follows:

Scheme A: Select the k BSUs with the largest demand, where k is the number of ASUs desired.

Scheme C: (1) Select the BSU with the largest demand as the first ASU.
(2) Assign all remaining BSUs to their nearest ASU.
(3) If the number of ASUs = k, Stop.
(4) Select the BSU with the greatest demand-weighted distance to its assigned ASU as the next ASU.
(5) Go to step (2).

Frequently planners want to analyze the trade-offs between the number of facilities required to cover the entire demand (or the maximal coverage attainable given a fixed number of facilities) versus the distance standard, S (for example, see Toregas and ReVelle 1973). A drawback to the implementation of our proposed rule 3 is that a new set of ASUs must be generated for every value of the covering distance, S, examined. It should be noted that an advantage of the aggregation schemes presented in Daskin et al. (1989) is that they are independent of the value of S, and consequently need to be calculated only once. However, given the relatively efficiency of implementing rule 3, this drawback does not appear severe.

SUMMARY

In this paper we have extended the concepts of demand data aggregation error to location problems involving covering. Due to the binary nature of the distance metric in such problems, the losses in locational information and resulting suboptimalities are potentially more significant than are similar errors in the p-median problem. Of the three sources of error described by Hillsman and Rhoda (1978), only Source C is not present in covering problems. Source A errors occur in two different cases: one where the facility covers the ASU but not the BSU assigned to it; and the other, where the BSU is covered by a facility but its associated ASU is not. The Source A Case 1 errors will always produce non-negative covering errors while Case 2 covering errors will always be nonpositive. The optimality errors that result from Case 1 may produce infeasible solutions in the
SCLP since not all BSUs may be covered. On the other hand, Case 2 errors may cause the SCLP to indicate that more facilities are needed than necessary.

Source B errors occur when a facility is sited at an ASU which has been assigned BSUs which are not covered by it. These errors produce the same type of inaccuracies (covering and optimality errors) as Source A Case 1 errors.

Daskin et al. (1989) have demonstrated that the effects of these errors may be significant in the MCLP, especially in their effect on the actual locations selected.

Two rules were proposed in this paper to reduce the effect of these errors when demand data are being aggregated at a set of ASUs. They suggest that (1) the ASUs should be a subset of the BSUs, and (2) BSUs should be assigned to an ASU only if they can be covered by a facility sited there. These rules will eliminate Source A, Case 2 and Source B errors in the SCLP and only Source B errors in the MCLP. A third aggregation rule was proposed which dictates the set of ASUs but eliminates both Source A and B error completely. Rule 3, which aggregates BSUs with identical covering sets, produces the greatest aggregation possible while still retaining all of the original locational information. While rule 3 is applicable to many different covering problems, the SCLP is better served by the specialized row reduction technique presented in Toregas and ReVelle (1973). However, their technique is based upon the assumption that all demand will be covered and consequently, cannot be implemented when this assumption is relaxed, for example, as it is in the maximal covering location problems.

Several computational tests were performed applying rule 3 to a 681-node location network. In these tests, demand data reductions of at least 19 percent were obtained. For practical covering distances (those used in the study employing the original data), the implementation of rule 3 reduced the problem from one-twentieth to two-thirds of its original size (for example, 681 demand nodes reduced to 31). These reductions were obtained with no loss of locational information, and consequently, no resulting covering or optimality errors. There were, however, sizeable reductions in solution time.

Clearly, the number of ASUs generated by rule 3 is a function of the underlying data. If the desired level of aggregation desired is not achieved, we recommend that the aggregation process start with rule 3 and then use scheme A or C proposed by Daskin et al. (1989).

LITERATURE CITED


