# Application of Geographic Information Systems in Planning Transit Services for People with Disabilities

MASSOUD JAVID, PRIANKA N. SENEVIRATNE, PRABHAKAR ATTALURI, AND JAY AGUILAR

The scarcity of data on the travel patterns and needs of people with disabilities makes the planning of public transit system to meet the Americans with Disabilities Act and other requirements a difficult process. This means that approximate methods have to be employed under very broad assumptions to estimate demand, select routes, and schedule services. Even the common approximations such as catchment area method can be cumbersome when route density is high and service area is large. The application of a geographic information system (GIS) to perform the necessary computations is described. It is shown that information such as block group population and percentages of people with disabilities, which is available from census records, and general travel characteristics of people with disabilities can be combined to arrive at reasonable estimates of demand for transit services. The use of the GIS for scheduling demand-responsive services where fixed-route service are unavailable is also demonstrated.

The area within a specific walking distance or travel time of bus stops is usually considered the catchment area or transit service area for purposes of estimating passenger demand. The demand is estimated as a function of the population within these service areas. This estimation process becomes increasingly complex as the length of the route, the number of stops, and catchment area population varies. When the population is composed of special groups such as people with disabilities, there is the added problem of dealing with very diverse needs.

Providers of on-call (demand-responsive) transit services for people with disabilities also face difficulties in scheduling and routing vehicles for pick-up and drop-off. If those functions can be optimized resources, they can be better utilized and the cost of operation can be lowered. Proper scheduling has also shown to enhance user perception of transit and thereby contribute to increased ridership.

The ability to easily estimate the expected number of users with disabilities in a transit service area can save transit planners and providers a great deal of time, particularly when evaluating alternatives route or planning new services. A simplified technique to estimate fleet size needed for providing on-call services and developing optimum schedules for existing and alternative services can also help the planners.

Geographic information systems (GIS) have the potential to become an ultimate time- and cost-saving tool of transit planners. Relevant information and data such as size of population with disabilities, origin and destination of trips, days and hours that trips are made, and so on can be incorporated into or form a GIS data base for corridor and transit service area analysis. Spatial analysis tools available in GIS can then be used to perform service area computations, scheduling, and routing.

This paper is based on a study that was undertaken to examine the pros and cons of using GIS for estimating demand for transit services by people with disabilities and scheduling demand-responsive transit vehicles. TransCAD Version 2.1 GIS package was chosen for this study. It is important to note that the focus and purpose of this study was neither to evaluate the available GIS software packages nor to indicate any preference for one software over the others. The GIS package was used to assess the ease with which computations can be performed and the challenges facing users of GIS when estimating demand for public transit.

## DATA BASE

Information for building the data base came from the Cache County 1990 Final Census Version of Topologically Integrated and Geographic Encoding and Referencing files (TIGER files, tgr49005.f41) and data from a questionnaire survey conducted as part of another study. First, the TIGER files were imported into TransCAD using its TCBuild program. County, state, census tracts, census blocks, block groups, intersections, and roads (including highways and streets) information was included in these files. Then, the 1990 Census data for the block groups were imported into the block group layer of this data base. These data included total population and disabled population by block group.

From this information, the following three data bases were created:

- Line data base (consisting of nodes and links),
- Area data base (consisting of census blocks, block groups, census tracts, county and state layers), and
- Point data base (with only one layer and nodes).

### Line Data Base

The line data base consists of a layer of nodes and a layer of links. Links represent road segments with nodes on either side, and nodes represent intersections and sometimes mid-block points

M. Javid, P. N. Seneviratne, and P. Attaluri, Department of Civil and Environmental Engineering, Utah State University, Logan, Utah 84322. J. Aguilar, Bear River Association of Governments, Logan, Utah 84321.

at the end of street segments. TIGER files were imported into the software using TIGER translate procedure in the TCBuild program. TIGER files contain information about road segments, and pipeline, rail road, landmark, and other topographical features. Road segment information was selected for the purpose of this study, and an ASCII file was created with that information. From the ASCII file, a line data base was built using the Build Data Base procedure in TCBuild. The data base stores all nodes and link data with respect to specific identification numbers assigned to them according to their longitude and latitude in space. The node layer consists of identification numbers (ID) and longitude and latitude columns. The links layer contains ID, longitude, latitude, name of the street segment, length of that segment, and an additional suffix to identify that segment. Sufficient empty fields were created while building the data base for further data manipulation. The line data base finally showed the road network throughout the county. Any calculations with respect to length can be performed with this data base. Figure 1 illustrates roads in the city of Logan and Cache County, Utah.

### Area Data Base

In the area data base, the census blocks in the study area were aggregated into 55 block groups, and these 55 block groups were further aggregated into 18 census tracts. All these layers (census blocks, block groups, and so on) contain the land area of each division. The area data base was also built in the same way as the line data base. Figures 2 and 3 show the total block group population for Cache County and the city of Logan, respectively.

### Point Data Base

The point data base contains information about points (e.g., major generators and attractors) such as addresses and coordinates. To build the point data base, an ASCII file with a few known spatial coordinates was created. These coordinates were drawn from the node layer of the line data base. Using the Build Data Base menu in TCBuild, the point data base was built for those points in the ASCII file. The ASCII file needs the ID number and longitude and latitude information in either fixed or comma-delimited format. The format of the input data and the columns in which each identifiable data is located in the input file has to be given using the Edit option before building the data base. More points can be added afterward in the point data base using the Geography Add procedure in the TransCAD program.

### Representation of Origins and Destinations

The trip origins and destinations of the people with disabilities surveyed during a previous study were input to a point data base. These points were represented by overlaying the labeled line data base on the point data base. This was done by simply using the same application file for both the data bases and making all the layers active layers. Figure 4 illustrates the primary origins and destinations of people with disabilities in Logan.

### Representation of Transit Routes

Separate data bases were built for each direction of the six fixed routes operated by the Logan Transit District (LTD). To do this,

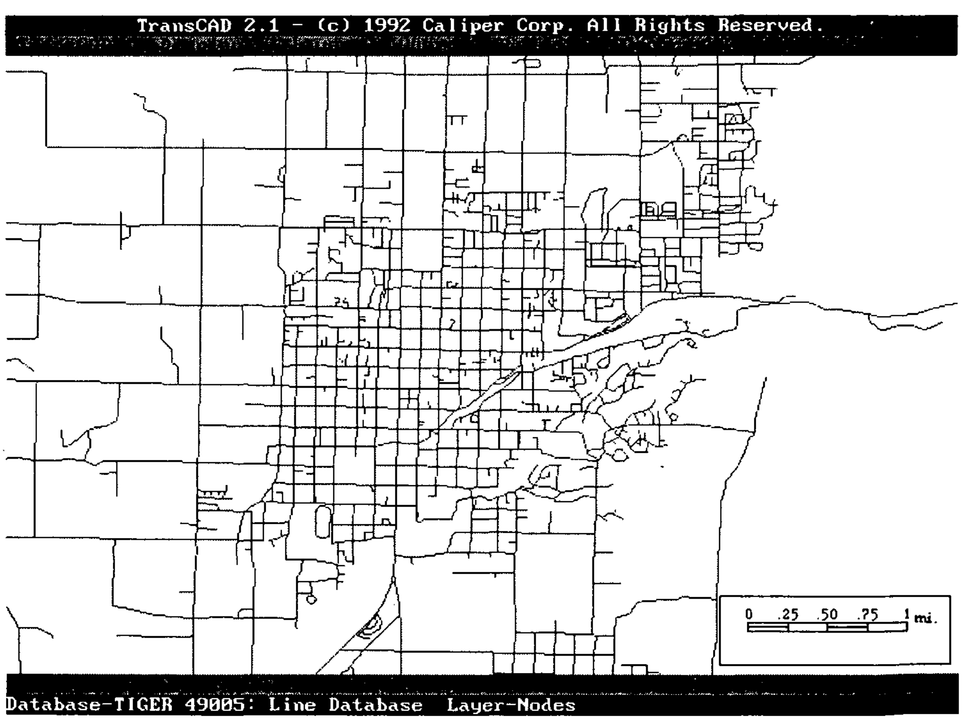


**FIGURE 1 Logan city road network.**

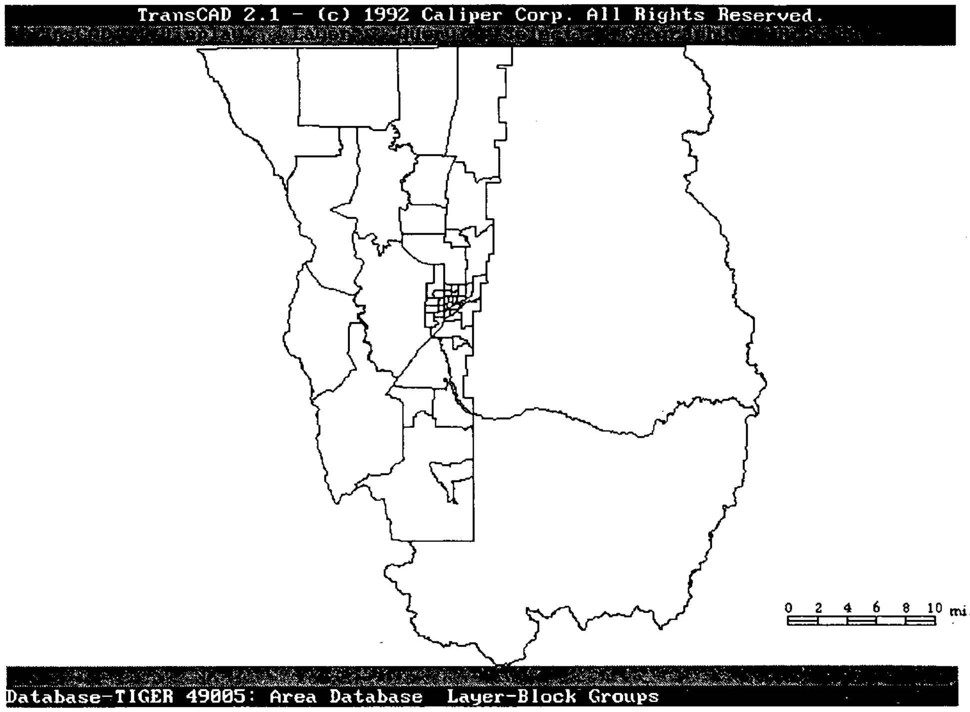


**FIGURE 2** Cache County block groups.

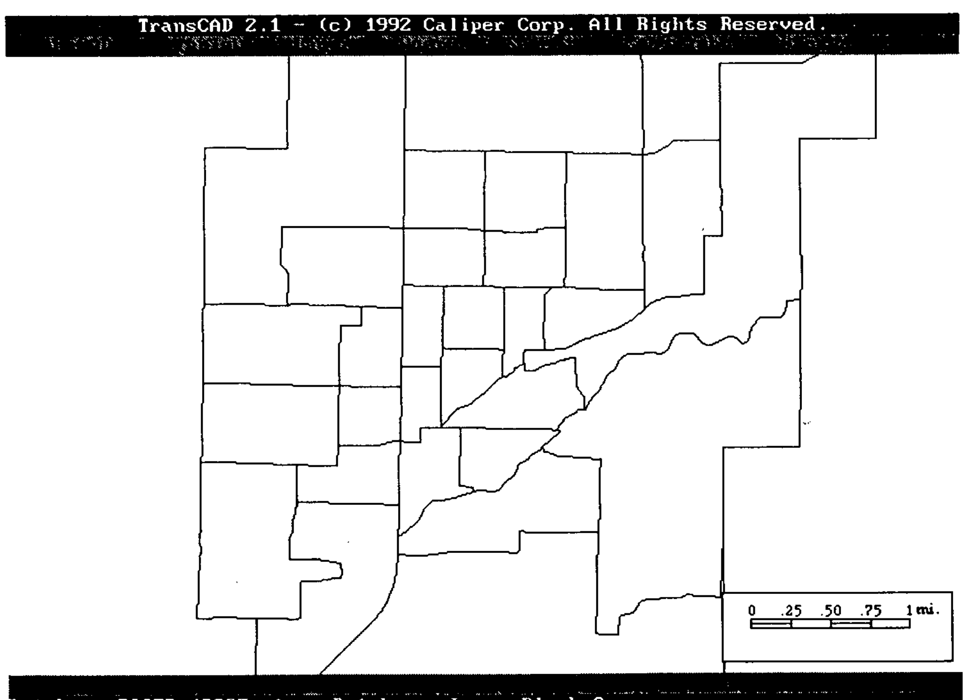


**FIGURE 3** Logan city road block groups.

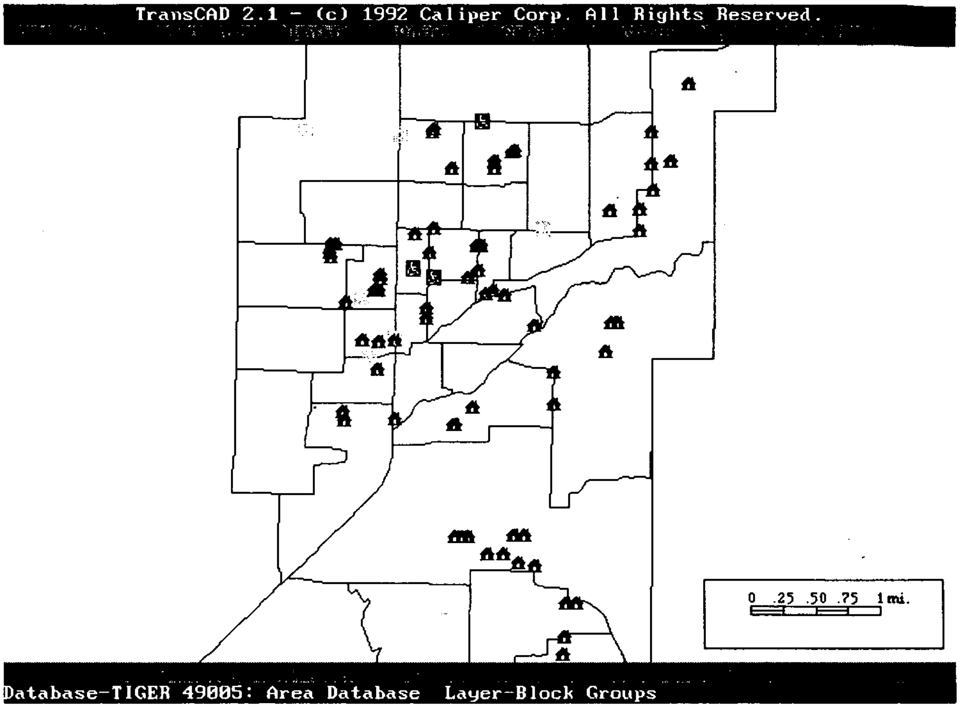


FIGURE 4 Origins and destinations of people with disabilities in Logan city.

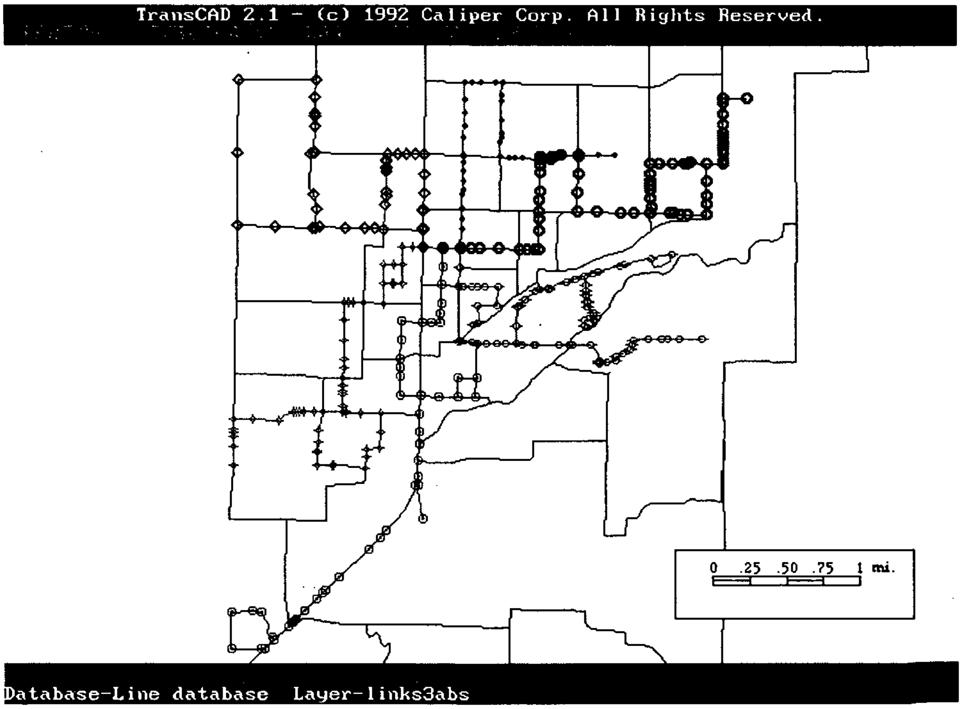


FIGURE 5 LTD transit routes in Logan city.

first the road layer was chosen as the current layer in TransCAD. Then the Select Several by Pointing option was used to select the streets that make up the transit route. The second step was to use the TCBuild program to export the selected links as an ASCII file and then use this file as a source to build a data base using TCBuild. This process was performed for each transit route separately. The LTD routes in Logan are shown in Figure 5.

### Definition of Demand

Two methods were used to estimate demand. But both methods are based on the assumption that demand is directly proportional to the number of people with disabilities (population) in the catchment area. Thus, first the service area had to be defined and then the population in the catchment area had to be estimated.

The catchment area for a given transit route has traditionally been defined as the area within a reasonable walking distance from the route or stop/terminals. A quarter mile, or approximately 400 m, has been considered in the past to be a reasonable walking distance (*1,2*).

In densely developed areas, when stop spacings are approximately 0.5 mi (800 m) or less, the catchment area can be considered to be 0.5-mi-wide corridor along the route. If stop spacings are larger, it has been suggested that a circular area with an approximately 400-m radius around each stop on the route (*3*) or various other forms and shapes of catchment areas be used.

In this case, it was decided to use the rectangular catchment area approach to estimate demand and compare it with another approach where demand is assumed to be proportional to the relative length of the transit route. With the latter approach, the population is assumed to be uniformly distributed along each road segment in an area; with the former approach, it is assumed to be distributed uniformly in the entire area.

### Estimation of Population

During the course of a study in Cache County (*4*), local service agencies had expressed concerns that the census data understated the number of people with disabilities in their service areas. These agencies have produced estimates considerably higher than the census. This discrepancy prompted a search for a valid estimate of people with disabilities in the study area.

The population of people with disabilities is most easily estimated as a proportion of the total population. In a previous study (*4*), it was estimated that the average percentage of people with disability in the study area is approximately 8 percent. Thus, in the present case, estimates of population in each block group were first obtained a follows:

$$p_i = 0.08P_i$$

where $p_i$ is the population of people with disabilities in block group $i$ and $P_i$ is the total population in block group $i$.

This approach has a shortcoming in that even if there are no people with disabilities living in a particular block group, 8 percent of the total population is assumed to have disabilities. To circumvent this problem, it was decided to use population figures given in the census data with an adjustment to account for the nonrepresentation of the younger population (under 16 years). The population was estimated in this case as

$$q_i = 2Q_i$$

where $q_i$ is the population of people with disabilities in block group $i$ and $Q_i$ is the census estimate of people with disabilities in block group $i$.

The multiplier (2) is the ratio of the Cache Valley Study (*4*) estimate to the census estimate of the percentage of people with disabilities.

### Estimation of Demand

#### *Area Method*

As described earlier, the demand under this method is assumed to be proportional to the ratio of the catchment area to total area in a given block group.

$$D = \sum_{\text{all } i} D_i = \sum p_i \text{ (or } q_i) \left(\frac{a_i}{A_i}\right) r_i$$

where

- $D_i$ = total demand for transit in block group $i$,
- $a_i$ = catchment area in block group $i$,
- $A_i$ = total area of block group $i$, and
- $r_i$ = probability that a person with a disability in block group $i$ uses transit.

#### *Road Segment (Linear) Method*

The demand is assumed to be proportional to the length of the road segments in the catchment area to the total length of all roads in the block group.

$$D = \sum_{\text{all } i} D_i = \sum p_i \text{ (or } q_i) \left(\frac{l_i}{L_i}\right) r_i$$

where $l_i$ is the length of road segments in catchment area and in block group $i$ and $L_i$ is the total length of all roads in block group $i$.

### Estimation of $r_i$

The probability or likelihood ($r_i$) of a person's using transit in block group $i$ depends on the composition of the people with disabilities in that block group. For instance, the Cache Valley study showed that people with mobility-related disabilities are more likely to use transit than those with hearing or visual impairments. Thus, $r_i$ is defined as

$$r_i = \sum_{\text{all } j} p\left(\frac{t}{d_j}\right) \cdot p_{d_j}$$

where $p(t/d_j)$ is the probability of a person with disability using transit given that the person has type $j$ disability, and $p_{d_j}$ is the proportion of people with type $j$ disability. $p(t/d_j)$ is available in some sources (*4*) where the population of people with disabilities has been identified and their transit use patterns have been examined.

**TABLE 1 Demand Estimations**

| Transit Route | Area Method | | Road Segment Method | |
|---|---|---|---|---|
| population | ($p_j$) | ($q_j$) | ($p_j$) | ($q_j$) |
| 1AB North | 44 | 37 | 44 | 36 |
| 1AB South | 43 | 70 | 39 | 61 |
| 2AB East | 70 | 51 | 59 | 37 |
| 2AB West | 34 | 41 | 32 | 34 |
| 3AB East | 51 | 53 | 52 | 50 |
| 3AB West | 40 | 61 | 45 | 70 |
| All Routes | 171 | 179 | 208 | 226 |

## Application of Demand Estimation Procedures

To illustrate the application of the procedures described, the demand for six fixed transit routes of the LTD were estimated separately and compared with one another.

The expected demand is given in Table 1 and based on two population estimates ($p_i$ and $q_i$) for the block groups discussed earlier. In both cases, a fixed $r_i$ of 8.4 percent, which is the estimated average percentage of person with disabilities using transit in Logan, was used to arrive at the demand estimates.

It can be seen that the method of estimating population does not produce significantly different results. However, the road segment method estimate of demand is approximately 24 percent higher than the area method.

## TransCAD Procedure Summary

### *Area Method*

To estimate the demand using the area method, a 0.5-mi buffer was created around each LTD transit route. To do this, the Query Buffer option in TransCAD was used. This option enables the user to select one or more segments, or the entire link layer. If several segments are to be buffered, the links must be selected before buffering using the Select option. The Buffer option will ask for the desired buffer width and the area layer to be buffered (in this case the block group layer). As soon as these two parameters are input, the results appear in a pop-up window, which can be printed. TransCAD Version 2.1 does not have the capability to save the results and the buffered area into a data base.

If there is the need to inspect, for example, an alternative transit route or an extension of a route for the increased coverage of people with disabilities, the link layer can be used to modify or create another layer or select desired segments for buffering. However, the buffer zones can overlap and result in double counting when the transit routes are spaced at less than the desired buffer width. In such cases, overlapping areas can be treated independently by creating a layer containing only the routes or the links that have overlapping service. If a portion of the block group area is inaccessible to the transit route because of a natural or manmade barrier or any other reasons, the buffering operation has no way of finding it or disregarding that portion in the analysis. This is a shortcoming of the area method, and the only way to overcome it is to calculate these areas manually.

### *Street Segment Method*

As mentioned previously, the street segment ratio method uses the ratio of the street segments of a block group in the analysis area to the total block group street length to estimate the population in the service area. To calculate the total street and road lengths in a block group, Data Editor Column Aggregation was used. Then the density of the people with disabilities along streets was

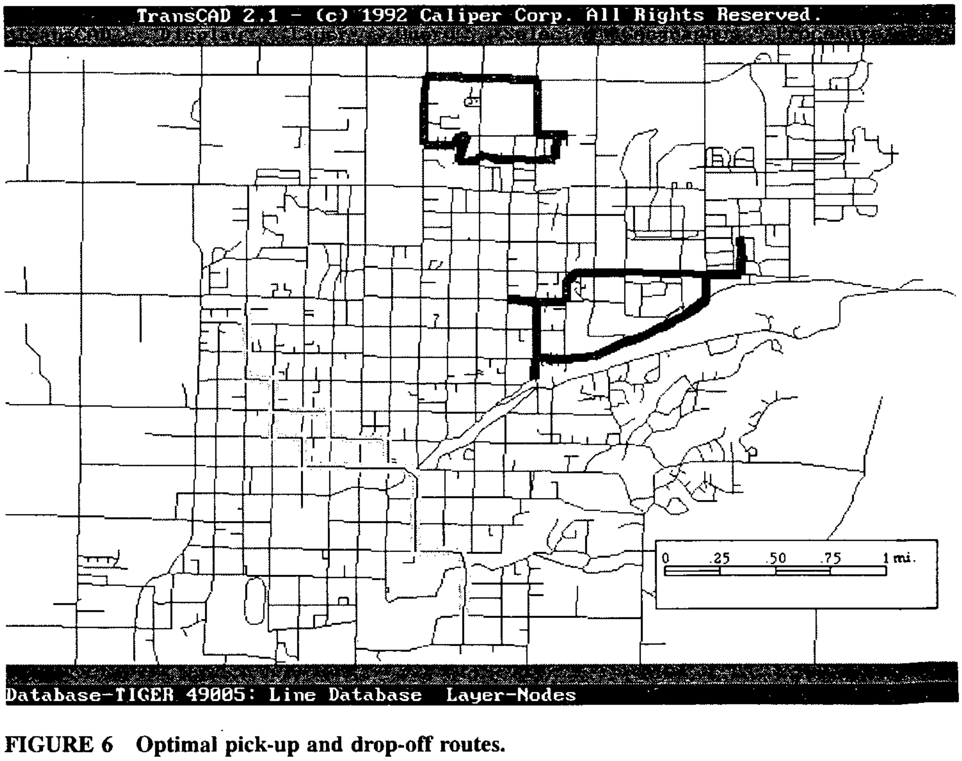


**FIGURE 6 Optimal pick-up and drop-off routes.**

```
              MVRF02 Report
              =============
Network File: C:\TRANSCAD\FINAL4.NET
Demand Table: C:\TRANSCAD\FINAL6.TAB
Shape Parameter: 1.0
Number of Trials: 20
Vehicle Capacity: 4.0

3 Depot(s):
       1850        3342        3611
10 Selected Points:
       2230        6909        2249        1894        3399
       3526        3776        3469        3276        3384
----------------------------------------------------------------------

[Depot]: 1850
trial         cost
    1         8.90
    2         8.90
    3         8.90
    4         8.90
    5         8.90
    6         8.90
    7         8.90
    8         8.90
    9         8.90
   10         8.90
   11         8.90
   12         8.90
   13         8.90
   14         8.90
   15         8.90
   16         8.90
   17         8.90
   18         8.90
   19         8.90
   20         8.90

[ROUTE PARTITION]:
Route 1: 4 customers:
 Customer  Node_id  Demand    Cost
     2       6909     1.0      2.5
     1       2230     1.0      2.2
     4       1894     1.0      1.6
     3       2249     1.0      1.0
 depot       1850              1.6
 Sum                  4.0      8.9

Total Cost:     8.9

[ROUTE SCHEDULE]: (*: customer, #: depot)

* Route 1 *
Customer Stop sequence: 0-2-1-4-3
Number of links on the route: 29.
   Node   Arrive       Node   Arrive       Node   Arrive
    1848 (  0:00)      2223 (  0:00)      2224 (  0:01)
    2398 (  0:01)      1846 (  0:01)      6909*(  0:01)
    2226 (  0:20)      2225 (  0:20)      2227 (  0:20)
    2236 (  0:20)      2239 (  0:20)      2234 (  0:21)
    2233 (  0:21)      2230*(  0:21)      1906 (  0:39)
    2229 (  0:39)      2245 (  0:40)      2244 (  0:40)
    1918 (  0:40)      1894*(  0:40)      1878 (  0:58)
    1880 (  0:58)      1881 (  0:59)      2252 (  0:59)
    2249*(  0:59)      2252 (  1:17)      2243 (  1:17)
    1879 (  1:17)      1850#(  1:18)
-----------------------------------------------------------------------

[Depot]: 3342
trial         cost
    1        12.95
    2        12.95
    3        12.95
    4        12.95
    5        12.95
    6        12.95
    7        12.95
    8        12.95
    9        12.95
   10        12.95
   11        12.95
   12        12.95
   13        12.95
   14        12.95
   15        12.95
   16        12.95
   17        12.95
   18        12.95
   19        12.95
   20        12.95
```

**FIGURE 7 Optimal routes and schedules** (*continued on next page*).

```
[ROUTE PARTITION]:

Route 1: 3 customers:
 Customer  Node_id  Demand    Cost
     2        3276     1.0     2.8
     1        3469     1.0     5.6
     3        3384     1.0     2.0
 depot        3342             2.6
 Sum                   3.0    12.9

Total Cost:    12.9

[ROUTE SCHEDULE]: (*: customer, #: depot)

* Route 1 *
Customer Stop sequence: 0-2-1-3
Number of links on the route: 42.
   Node   Arrive       Node   Arrive       Node   Arrive
    3344 (  0:00)       3345 (  0:00)       3347 (  0:01)
    3348 (  0:01)       3349 (  0:01)       3311 (  0:01)
    3276*(  0:02)       3311 (  0:20)       3349 (  0:20)
    3348 (  0:20)       3347 (  0:21)       3392 (  0:21)
    3401 (  0:21)       3426 (  0:22)       3453 (  0:22)
    3452 (  0:23)       3458 (  0:23)       3459 (  0:23)
    6947 (  0:23)       6973 (  0:23)       6949 (  0:23)
    6950 (  0:23)       3456 (  0:23)       3470 (  0:24)
    3469*(  0:24)       3470 (  0:42)       3456 (  0:42)
    3419 (  0:42)       3388 (  0:43)       3389 (  0:43)
    3384*(  0:43)       3389 (  1:02)       3388 (  1:02)
    6967 (  1:02)       6966 (  1:02)       3395 (  1:02)
    3394 (  1:02)       3353 (  1:02)       3339 (  1:03)
    3340 (  1:03)       3341 (  1:03)       3342#(  1:03)
-----------------------------------------------------------------------

[Depot]: 3611
trial        cost
    1       19.01
    2       19.01
    3       19.01
    4       19.01
    5       19.01
    6       19.01
    7       19.01
    8       19.01
    9       19.01
   10       19.01
   11       19.01
   12       19.01
   13       19.01
   14       19.01
   15       19.01
   16       19.01
   17       19.01
   18       19.01
   19       19.01
   20       19.01

[ROUTE PARTITION]:
Route 1: 3 customers:
 Customer  Node_id  Demand    Cost
     3        3776     1.0     4.5
     2        3526     1.0     7.3
     1        3399     1.0     2.2
 depot        3611             5.0
 Sum                   3.0    19.0

Total Cost:    19.0

[ROUTE SCHEDULE]: (*: customer, #: depot)
* Route 1 *
Customer Stop sequence: 0-1-2-3
Number of links on the route: 59.
   Node   Arrive       Node   Arrive       Node   Arrive
    3609 (  0:00)       3610 (  0:00)       3605 (  0:00)
    3571 (  0:01)       3569 (  0:01)       3524 (  0:01)
    3526 (  0:02)       3471 (  0:02)       3468 (  0:02)
    3432 (  0:03)       3433 (  0:03)       3405 (  0:03)
    3399*(  0:03)       3405 (  0:21)       3433 (  0:22)
    3432 (  0:22)       3468 (  0:22)       3471 (  0:23)
    3526*(  0:23)       3524 (  0:41)       3527 (  0:41)
    3534 (  0:42)       3533 (  0:42)       3572 (  0:42)
    3574 (  0:43)       3575 (  0:43)       3576 (  0:43)
    7018 (  0:43)       3613 (  0:43)       3612 (  0:43)
    3614 (  0:44)       4310 (  0:44)       3648 (  0:44)
    3647 (  0:44)       3658 (  0:44)       3689 (  0:44)
    3710 (  0:45)       3714 (  0:45)       3720 (  0:46)
    3728 (  0:46)       3673 (  0:46)       3757 (  0:46)
    3765 (  0:46)       3776*(  0:46)       3765 (  1:05)
    3757 (  1:05)       3673 (  1:05)       3728 (  1:05)
    3720 (  1:05)       3714 (  1:06)       3710 (  1:06)
    3689 (  1:06)       3658 (  1:07)       3647 (  1:07)
    3648 (  1:07)       4310 (  1:07)       3614 (  1:07)
    3612 (  1:07)       3611#(  1:08)
-----------------------------------------------------------------------
```

**FIGURE 7** *(continued)*

determined for each block group. The next step was to transfer data from the block group layer to road layer, in which the Data Editor Tag function was used. This was done for the estimates of people with disabilities using the first and second methods.

Conventional buffering of the street segments along the transit route would result in inclusion of the street segments that are within the buffered area but their walking distance is more than the desired value or they are not connected to the streets with transit service due to the existence of a barrier. To remedy this, the Arc/Node partitioning capacity of TransCAD was used. Nodes of each transit layer were used to partition the road segments that were connected to the transit route and were within the 0.5-mi buffer area. This operation selected all the segments within the service area including the segments that were partially in the service area. To find the portion of length of street segments in the analysis area, a simple conditional algorithm was used in Data Editor, and the resulting length was multiplied by the disabled population density of the road segment. The selected links can be exported into a spreadsheet program (e.g., Lotus 1-2-3) to sum the columns for the disabled population. This operation was performed for each transit route in the present case.

## VEHICLE ROUTING AND SCHEDULING

The origin and destination information obtained from a previous survey (*4*) was used to create a hypothetical routing assignment. To do this, 10 origins were selected as the pick-up addresses and three destinations were chosen as drop-off locations. Then a road network covering all the pick-up and drop-off locations was created using Build Network. Next, multiple depot vehicle routing assignment was performed using Procedure Routing and Scheduling Models-MVRP02. This procedure identifies a pick-up/drop-off strategy for vehicles by minimizing costs. Cost can include a variety of components such as travel time and operating cost. In this case the travel time (based on an average 20-mph operating speed) was chosen to be minimized, and one drop-off point was allocated to each pick-up point. For this example the following parameters were used:

| *Parameter* | *Value* |
|---|---|
| Pick-up point | 10 |
| Drop-off point | 3 |
| Demand at each pick-up point | 1 passenger |
| Vehicle capacity | 4 passengers |
| Fixed service time at each pick-up/drop-off point | 0.1 hr |
| Variable service time at each pick-up/drop-off point | 0.5 hr |

After 20 iterations, it was found that three vehicles were needed to minimize travel time between the chosen origins and destinations. The output included the routing and arrival times at each pick-up and drop-off point on the route. Figure 6 shows the three optimum routes, and Figure 7 illustrates the fleet size requirement and the pick-up strategy for each route.

It should be noted that the solution does not include the travel path to/from depot. Nevertheless, this analysis can be used to identify the cluster of origins with common destination and their optimum route. It can also be used to find an optimum location for vehicle depot(s) with respect to optimum routes of these clusters.

## CONCLUSIONS

The demand for each of the six LTD fixed routes and the entire network was estimated using two methods under two different estimates of people with disabilities. Each method of estimation has its own merits and demerits. One method could be more suitable for a particular set of block groups whereas the other might be more suited for another set. In the present case, because of the existence of many parcels of agricultural land use in the study area, the road segment method derived better estimates. The software was also better suited for treating inaccessible road segments than inaccessible area portions, which must be excluded when using the area method.

Even though a GIS can perform the computations and display results within a matter of seconds, acquisition of suitable information and the development of data bases are major challenges. Most basic information is available from TIGER files and census records, but examining travel patterns, estimating the propensity to use transit ($r_i$), and refining the data bases will continue to be time-consuming and expensive tasks. For instance, the road layer for Cache County extracted from TIGER files did not have address information, which reduced the precision of vehicle routing and scheduling. The remedy to this problem, which would involve manually gathering starting and ending address numbers of each street segment either by physical inspection or by obtaining the information from the city planning office, would be prohibitively expensive. Entering this information into the proper data base for each road segment would also consume several days.

In summary, a GIS can be a valuable tool for estimating demand and developing strategies to satisfy the demand by proper routing to gain maximum demand coverage. However, the results or planning information generated by GIS depends heavily on the precision of input data, and the benefits will only be derived in the long run.

## ACKNOWLEDGMENTS

Funding for the study on which this paper is based was provided by Project ACTION and the Mountain Plains Consortium.

## REFERENCES

1. *Canadian Transit Handbook*, 2nd ed. Canadian Urban Transit Association and Road and Transportation Association of Canada, Ottawa, Ontario, 1985.
2. Pagano, A. M., and C. E. McKnight. Quality of Service in Special Service Paratransit: The Users' Perspective. In *Transportation Research Record 934*, TRB, National Research Council, Washington, D.C., 1983, pp. 14–23.
3. O'Neill, W. A., R. D. Ramsey, and J. Chou. Analysis of Transit Service Area Using Geographic Information Systems. In *Transportation Research Record 1364*, TRB, National Research Council, Washington, D.C., 1992, pp. 131–138.
4. *The Transportation Needs of People with Disabilities in Cache County.* Project ACTION Final Report. National Easter Seal Society, Washington, D.C., 1993.

*Publication of this paper sponsored by Committee on Transportation Data and Information Systems.*