Data Requirements and Data Sources for Transit GIS Applications

National Urban Transit Institute

at the University of South Florida's Center for Urban Transportation Research

A consortium with • Florida State University • Florida A & M University • Florida International University

Urban Transit Institute

CENTER FOR URBAN TRANSPORTATION RESEARCH
College of Engineering, University of South Florida
4202 E. Fowler Avenue, CUT 100, Tampa, FL 33620-5375
(813) 974-3120, SunCom 574-3120, Fax (813) 974-5168
Gary L. Brosch, CUTR Director
Steven E. Polzin, NUTI Director

Consortium Member Contacts

UNIVERSITY OF SOUTH FLORIDA
Steven E. Polzin, PhD, PE
Center for Urban Transportation Research
University of South Florida
4202 E. Fowler Avenue, CUT 100, Tampa, FL 33620-5375
(813) 974-9849, Fax (813) 974-5168
E-mail: *polzin@cutr.eng.usf.edu*

FLORIDA STATE UNIVERSITY
Bill Mustard
Director of Research and Services
The Marketing Institute
Florida State University
College of Business, Tallahassee, FL 32306-1111
(850) 644-2509, SunCom 285-2509, Fax (850) 644-6231
E-mail: *bill@fimat.cob.fsu.edu*

FLORIDA A&M UNIVERSITY
Charles A. Wright, PhD, PE, Professor
College of Engineering Sciences, Technology & Agriculture
Florida A&M University
Engineering Technology, Room 105
Post Office Box 164, Tallahassee, FL 32307
(850) 561-2920, SunCom 286-3506, Fax (850) 561-2248

FLORIDA INTERNATIONAL UNIVERSITY
L. David Shen, PhD, PE
Chairman/Director, Lehman Center for Transportation Research
Department of Civil and Environmental Engineering
Florida International University
10555 W. Flagler St., Rm 3685, Miami, FL 33174
(305) 348-3055, SunCom 441-3055, Fax (305) 348-2802
E-mail: *shen@eng.fiu.edu*

National Urban Transit Institute

at the CENTER FOR URBAN TRANSPORTATION RESEARCH
University of South Florida
Florida State University
Florida A&M University
Florida International University

Data Requirements and Data Sources for Transit GIS Applications

Fang Zhao and L. David Shen
Principal Investigators
Florida International University

February 1998

Lehman Center for Transportation Research
College of Engineering and Design
Florida International University
The State University of Florida at Miami
Miami, FL 33199
TEL: (305) 348-4103
FAX: (305) 348-4057

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No. NUTI96FIU1.3	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Data Requirements and Data Sources for Transit GIS Applications		5. Report Date February 1998
		6. Performing Organization Code
7. Author(s) Fang Zhao and L. David Shen		8. Performing Organization Report No.
9. Performing Organization Name and Address National Urban Transit Institute Lehman Center for Transportation Research, Florida International University Park, Miami, FL 33199		10. Work Unit No.
		11. Contract or Grant No. 21-17-230-L0-B
12. Sponsoring Agency Name and Address Office of Research and Special Programs U.S. Department of Transportation, Washington, D.C. 20690		13. Type of Report and Period Covered Sept. 1996 through December 1997
		14. Sponsoring Agency Code

15. Supplementary Notes

Supported by a grant from the U.S. Department of Transportation, University Research Institute Program

16. Abstract

Geographic Information Systems (GIS) have been recognized as invaluable information management and decision making tools for public transit planning and management. Many transit properties have made various levels of effort to use GIS in transit planning and operations. These efforts, however, are sometimes motivated by immediate needs and are often limited in their sophistication. There are many challenges to be faced if GIS is to be used in transit widely and effectively. Some of the major challenges, as identified in (Schweiger 1992), include funding, resources and training, lack of recognition of GIS capabilities or the value of GIS, and data issues. GIS data issues have not been adequately addressed and transit properties, especially small properties without resources to hire GIS professionals, are in need of information on GIS applications and existing data sources to be able to implement GIS applications. This research involves the identification of the data needs for various transit GIS applications and possible sources from which useful GIS data may be obtained. Possible GIS transit applications, the data needs for these applications, the existing and potential data sources, and other issues related to data quality and maintenance are discussed.

17. Key Words GIS, public transit, data, database, data sources	18. Distribution Statement Available to the public through the National Technical Information Service (NTIS), 5285 Port Royal Road, Springfield, VA 22181, ph (703) 487-4650		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of pages	22. Price

Form DOT F 1700.7 (8-69)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

In the course of this project, we have received kind assistance from many people from different transit and government agencies. We would like to express our sincere thanks to the following individuals for their assistance:

Velma Ceballos, System Analyst, Information Technology Division, Miami-Dade County, FL.

Lina C. Kulikowski, Associate Planner, GIS, Planning Information Technology Division, Broward County, FL.

Spear Bruce, Assistant Director, Bureau of Transportation Statistics, U.S. Department of Transportation, Washington, D.C.

We would also like to thank Ms. Qing Xia, a graduate research assistant at the Lehman Center for Transportation Research, for making important contributions to the project. Ms. Vivian Donis' work as the technical editor of this report is also greatly appreciated.

This project is made possible through a grant from the U.S. Department of Transportation, Research and Special Programs Administration. Their support is gratefully acknowledged. The views expressed here are those of the authors' and do not reflect the opinions or policies of the U.S. Department of Transportation.

ABSTRACT

Geographic Information Systems (GIS) have been recognized as invaluable information management and decision making tools for public transit planning and management. Many transit properties have made various levels of effort to use GIS in transit planning and operations. These efforts, however, are sometimes motivated by immediate needs and are often limited in their sophistication. There are many challenges to be faced if GIS is to be used in transit widely and effectively. Some of the major challenges, as identified in (Schweiger 1992), include funding, resources and training, lack of recognition of GIS capabilities or the value of GIS, and data issues. GIS data issues have not been adequately addressed and transit properties, especially small properties without resources to hire GIS professionals, are in need of information on GIS applications and existing data sources to be able to implement GIS applications. This research involves the identification of the data needs for various transit GIS applications and possible sources from which useful GIS data may be obtained. Possible GIS transit applications, the data needs for these applications, the existing and potential data sources, and other issues related to data quality and maintenance are discussed.

EXECUTIVE SUMMARY

INTRODUCTION

Geographic Information Systems (GIS) have been recognized as invaluable information management and decision making tools for public transit planning and management. Many transit properties have made various levels of effort to use GIS in transit planning and operations. These efforts, however, are sometimes motivated by immediate needs and are often limited in their sophistication. There are many challenges to be faced if GIS is to be used in transit widely and effectively. Some of the major challenges, as identified in (Schweiger 1992), include funding, resources and training, lack of recognition of GIS capabilities or the value of GIS, and data issues.

Data are a critical element in GIS. GIS is data intensive, not only in terms of the tremendous volumes of data used, but also in terms of resources required for data collection and conversion, which often amount up to 80 percent of the total GIS investment costs. While there have been studies in the past on transit information management systems (TRB 1994, Hill 1995) and the state-of-the-art transit GIS applications (Schweiger 1992, Moyer 1995, CUTR 1996), GIS data issues have not been adequately addressed. Existing literature on GIS standards (e.g. Racca 1995, FTA 1996) is more data- or database-oriented than application-oriented, addressing issues mainly on file formats, data exchange formats, and database design. Transit properties, especially small properties without resources to hire GIS professionals, however, are in need of information on GIS applications and existing data sources to be able to implement GIS applications. Providing such information is therefore the goal of this research.

RESEARCH OBJECTIVES AND METHODOLOGY

This research is intended to identify the data needs for various transit GIS applications and possible sources from which useful GIS data may be obtained. The objective is to help transit properties, especially small and medium transit properties which often lack the technical expertise and resources required by sophisticated GIS applications, better plan GIS implementations by providing information on possible GIS transit applications, the data needs for these applications, the existing and potential data sources, and other issues related to data quality and maintenance. This report, however, may also be used by transit professionals and GIS analysts as a reference for transit GIS applications and GIS data sources.

The research involved a comprehensive literature review. The purpose of this review is to identify the current GIS applications in transit, the data used, and the methods by which the data are collected.

To systematically identify the data needs for transit GIS applications, known and potential GIS transit applications related to transit planning, operations, and management are studied. Based on the specific applications or types of applications, the data needs and possible data sources are investigated, including those that may become available with the implementation of an advanced transit technology such as AVL and APC. The general characteristics of different data sources including the contents, accuracy, frequency of change, currency, maintenance, etc., are studied.

CURRENT TRANSIT GIS APPLICATIONS

GIS applications have been implemented at many transit agencies over the past 10 years. A 1992 study by Schweiger provides an overview of these GIS applications in public transit nationwide. Schweiger conducted 74 interviews with 67 organizations across 30 states, including 46 transit agencies, and 21 Metropolitan Planning Organizations (MPOs). The results of that study showed that GIS was being used or being implemented at that time for a wide variety of applications. Among the 46 transit agencies, 21 have GIS, and of the 21 MPOs, 15 have GIS. In 1995, another survey was conducted by Harman (1995), 269 entities were contacted, and 202 survey forms were completed. Later, all transit agencies and MPOs in urban areas with a population of over 200,000 were interviewed. Nearly every transit agency (TA) with a fleet of 500 and above indicated that they were using GIS in some form or another. GIS is used in transit agencies mostly in areas related to transit planning, transit operations, such as scheduling and run cutting, marketing, fixed route and paratransit dispatching, customer services, and facility management. MPO GIS uses are mainly for forecasting ridership, service planning, and development of map products.

DATA REQUIREMENTS FOR TRANSIT GIS APPLICATIONS

Data needs for different transit GIS applications are described in this section, which are organized by major transit function areas. The transit functions, possible GIS applications, and data needed for such applications are described.

Transit Ridership Forecast

GIS may be used in transit ridership forecast (or long term transit planning) in the following ways:

- Network building, data creation and extraction for transportation analysis zones, and spatial aggregation of data.
- Analysis of route level demographics, such as displaying population living within a certain distance of a particular transit route.

- Modeling of proposed routes or extensions to existing routes to determine projected ridership, study of origins and destinations of home-based work trips, and generation of listings of street names and address ranges that fall within the study area.

- Identification of areas of high population density and low income and the proportion of families that do not own a car. By comparing these parameters with the transit service provided, areas that need improved transit services may be determined.

The core data in transit ridership forecasts are demographic and transit data, which are described below.

- Household. The number and type of dwelling units (low-density single family, medium-density single family, multi-family units, apartments, etc.) may be used to further determine distribution of population within a geographic analysis unit.

- Family Income. Total amount of income of persons 15 years old and over in each family. The relationship of income and travel by transit is often reversely proportioned.
- Car ownership. The number of cars owned by each family is also reversely proportional to transit use.

- Population density. Population density represents the number of people residing per unit area. It best describes the daily trip potential in the origin area.

- Employment or employment density. Employment density represents the number of jobs per unit area. Data may be obtained from local MPOs. More accurate data may be available commercially regarding the actual locations of businesses and number of employees working at the location.

- Trip attractors. Trip attractors include types and locations of employment, retail and recreational facilities, hospitals and medical centers, schools, public libraries, etc.

- Transit data. These include street network and transit routes, transit service frequency and coverage, travel times, and transit trip costs.

Service Planning/Market Analysis

Some major GIS applications for service planning and market analysis may include the following:

- Display of the relationship of a new route to existing routes.

- New route design. GIS can help control the duplication of bus routes when designing a new route.

- Stimulation of stop spacing, siting, and travel time for newly designed routes.

- Level-of-service monitoring.

- Examination of sub-area bus circulation and terminal locations.

A potential service planning instrument that may be useful is the regional household survey data collected periodically for urban transportation planning and modeling. Using GIS, the survey data may be analyzed to identify deficiencies in service configuration in terms of disparity between travel pattern and service route configuration. For a similar application, GIS may be used to assess mobility problems for welfare recipients. By storing addresses of welfare recipients and addresses of potential employers, transit services may be analyzed to evaluate whether these potential employers are accessible to welfare recipients. In addition, accessibility to low cost child care and job training facilities should also be provided.

Key data for service planning include:

- bus stop, route and running time data;

- the density and distribution of income and age structure of population;

- the available street pattern with particular regard to the suitability of the bus service, including street segments, intersections, etc.;

- locations of employment, retail and recreational areas, hospitals and medical centers, child care facilities, schools, and job training sites;

- demographics and socioeconomic data including welfare recipients' addresses.

Service/Operations Monitoring

Three important operational measures may be examined by using GIS:

- Schedule adherence may be monitored for the same or for different routes for schedule adjustments or evaluating the driver's performance.

- Passenger loads as well as the change in the load pattern may be monitored over time.
- Fare revenue data collected for each route may be analyzed to evaluate route performance.

With the introduction of new technology such as automatic vehicle location (AVL) and automatic passenger count (APC), running time and ridership data may be collected at any time, and buses will be continuously monitored for schedule adherence and emergency situations. If an agency has implemented APC, it may utilize the data such as the number of seats on a bus and the maximum passenger capacity of transit vehicles to determine the suitable fleet size for each route.

Data used in operations monitoring include (Elbadrawi 1996):

- Driver performance
- Vehicle schedule adherence
- Passenger loads
- In-vehicle and off-vehicle security/safety
- Local traffic conditions

Customer Information Services

GIS may be used as a powerful tool to provide the transit users with information about transit services on a digital map that may be used for pre-trip, in-terminal, or in-vehicle planning. Travel information may be delivered interactively via telephone, kiosks at public places, and the Internet. Functions of customer information system include:

- Monitoring and responding to passenger comments and complaints;
- Disseminating routes and schedule information;
- Disseminating real-time vehicle and route status;
- Disseminating fare information;
- Receiving incoming service requests (for demand responsive service); and
- Disseminating ADA-required mobility information.

The data needed includes transit network, transit routes, schedule or fare information, departure time, real-time vehicle data, etc.

Facility and Real Estate Management

Transit infrastructure inventory data are crucial in facility management. A GIS may be employed to provide better management of spatially distributed infrastructure data. The main contribution of GIS is its ability to geographically reference the facilities to routes, street segments, and to actual geographic coordinates.

There are three kinds of facilities in transit organizations that involve the geographical referencing of inventories of facilities (CUTR 1996):

- major facilities like bus maintenance and dispatch center facilities that have street addresses;
- bus shelters and bus signs that exist in street rights-of-way that do not have street addresses; and
- separate rights-of way for rapid transit facilities.

The attribute data needed for each inventory of facility based on previous analyses conducted by FTA are listed below (Richard 1997):

- Age: Asset age measured in years. Increasing age implies declining physical condition.
- Condition: This connects with the facility usage. An increasing rate of usage will increase the decay rate.
- Maintenance rate: An increasing rate of maintenance will slow down deterioration. Measurements of the rate of maintenance vary with facility type.

These data should be used to develop forecasting models and maintenance strategies and programs.

Transit Safety and Security

GIS may be employed in a transit safety improvement program by allowing analysis of accident data in response to complaints related to individual sites (Mark *et al*, 1990) and identification of unsafe factors. For example, by combining the location and operation data of bus route, operation procedures may be checked to avoid dangerous corners or intersections, provide adequate bus-stop distances for boarding, alighting, and merging with traffic, avoid railroad crossings, and adhere to traffic controls.

The data needed for transit safety analysis include vehicle operating data including headway, schedule, bus trip, etc., facility maintenance data, core transit data such as bus stops and route locations, street network, and accident records.

ADA and Paratransit

GIS may be used to maintain a database of paratransit users, study demand for paratransit services, analyze paratransit users' needs for accessibility to employment, medical, and social services, and assist in scheduling and dispatching of paratransit vehicles. Paratransit users may be address-geocoded in order to maintain a database of customers. This database is useful to locate a paratransit user by retrieving all the relevant information of address, location, and special transportation needs of the passenger. For individual trips, GIS may be used for vehicle routing to determine trip assignments for each vehicle. Records of services including origin, destination, travel time, travel frequencies, etc., may be useful in analyzing travel patterns of paratransit users to determine the best strategies to meet the demand.

Data needed for this application include customer addresses, trip attraction locations, street network, impendence on street network, transit network, and related service information.

TRANSIT GIS DATA SOURCES

Data Analysis Unit

For the purpose of aggregating data, transportation professionals need to process data by geographic location. Two kinds of data units are commonly used. One is census blocks or tracks, and the other is traffic analysis zones (TAZ), the former of which are data units used by U.S. Bureau of Census for collecting and compiling census statistics, and the latter is often used by transportation professionals.

Criteria for Evaluating the Quality of a Data Source

A data source may be evaluated by considering its accuracy, completeness, timeliness, correctness, credibility, reliability, convenience, precedence, and maintainability (Montgomery, 1993).

Existing Public and Private Data Sources

Public data sources refer to government agencies that produce GIS products. Data from public sources are most commonly used by transit agencies. It also forms the basis of various data products provided by private firms. The cost for public sources is mostly only to recover staff expenses and materials, therefore public data sources represent a great resource for low-cost data acquisition.

Public Data Sources

TIGER/Line File

TIGER/Line file contains the following geographic features and some aatributes for these features:

- Line features: roads, railroads, hydrography, miscellaneous transportation features and selected power and pipe lines;

- Landmark: point landmarks such as schools and churches, area landmarks such as parks and cemeteries, Key Geographic Location (KGL) such as a building;

- Polygon: geographic entity codes for areas used to tabulate the 1990 census statistical data and current geographic areas, locations of area landmarks.

USGS Digital Line Graph File

The DLG contains the cartographic information extracted from 1:100,000 and 1:24,000 map sheets for the United States, which are developed through orthophotos. The most recent set of photos represent the year of 1994. These maps are extremely good in accuracy and quality, particularly the 1:100,000 sheets, which now cover the entire United States. DLG information can be directly pipelined into a GIS using DLG as the exchange format. DLG files provide the following data:

- Boundaries - Includes political and administrative boundaries.

- Hydrography - Includes flowing water, water bodies, wetlands, and related features.

- Manmade Features - Includes built-up areas and populated places. State capitals and county seats are also identified.

- Transportation - Includes major transportation systems.

- U.S. Public Land Survey System -- Includes land grants and subdivisions of the public lands to the township and range level.

The deficiency of USGS DLG files lies in that it doesn't include address information or land use type. The best way to assign addresses to these data would be to use TIGER/Line files and match them with streets identified by DLG files.

U.S. Census Statistics

The 1990 census data products, being released during 1991-93, are available in a variety of new and traditional media. Summary Tape Files (STF's) are database files providing statistics with greater subject matter with more details than printed reports. They also present statistics for geographic units such as block groups and blocks, which are not included in the reports. STF 1 includes 100 percent population and housing counts and characteristics down to the block level. STF 2 contains 100 percent population and housing characteristics, showing more subject details than STF 1. STF 3 includes sample population and housing characteristics down to the block level.

Census Transportation Planning Package (CTPP)

The Census Transportation Planning Package (CTPP) is a set of special tabulations of 1990 census data tailored to small geographic areas, traffic analysis zones (TAZs), to meet the data needs of transportation professionals. CTPP is developed from the long form of the Decennial Census. The data are derived from a 17 percent sample and are updated every 10 years.

The tables contain information of population and household characteristics, worker characteristics and characteristics of Journey-to-work (JTW). Similarities exist between the CTPP and other Census products, such as STF (Summary Tape File) 3A (JHK & Associate, 1995). Notable differences also exist. The CTPP is the only Census product that summarizes tables by both place of work and by place of residence while others provide only information by place of residence.

National Transit GIS Database

The Federal Transit Administration National Transit Geographic Information System (NTG) is a representative inventory of the public transit assets of the country. By the Fall of 1996, 550 of the nations' fixed route bus services had been built into the GIS route systems.
The minimum data products for each transit agency included in the database are as follows:

- A GIS line database of the streets of the county or counties where the service is located. The street data are extracted from the Bureau of Census 1992 Enhanced TIGER files.
- A condensed file of the TIGER-based street network that takes up less data storage and provides quicker access and display.
- A GIS network built from the county street database.

- A GIS route system of the bus routes, built from selected TIGER street segments as indicated on the transit agency's route maps and schedules.

Local Government Data Sources

There are a variety of data sources in local government agencies. These data are developed for specific purposes and can be used as a land base or for trip generation purposes in transit GIS applications. The typical data sources include the following (Montgomery et al, 1993):

- Community oriented data such as property address files, building permit files, demographic information, zoning files, accident and crime statistics, environmental conditions, and hazardous materials information (Glenn, 1993).

- Assessment oriented data including property tax records. Information can include business, square footage, and number of employees or owned assets.

- Land use data such as those compiled by state water management districts. They can provide land use classification information, such as an urban and built-up area, agricultural range, water, transportation, communications, and utilities data.

- Accounting oriented data.

Private Sources

There are numerous companies that provide commercial geographic data. Here we briefly introduce some well known data providers that provide demographic, street center line, and boundary file data of the United States. **Table S.1** summarizes the data provided by major private data providers.

Table S.1 Summary of Data Provided by Major Private Data Providers

ETAK, Inc.	road network coverage, addresses, ZIP codes, population, road classification, political and statistical geography, major landmark point features in selected areas, urban one-way street encoding in selected areas
Navigation Technologies	road network geometry, connectivity of roadway, roadway names, block-by-block address ranges, postal codes and political entities, vanity address handling (such as IBM Plaza, United Nations Plaza), roadway classifications, roadway characteristics, overpasses and underpasses
America.dbf, Inc.	census demographics with transportation attributes, national street and boundary digital map set
Claritas Corporation	current-year estimates and five-year projections for census variables,
Business Location Research	street network, traffic volumes, major roads, highway network, demographic data
Dun & Bradstreet	businesses and their attributes including name of the firm, address, latitude/ longitude, number of employees by location, by SIC code and a contact person
American Business Information, Inc.	business name, location, mailing address, type of business, number of employees, professional specialities, and latitude and longitude coordinates

Date Conversion Ability in Popular GIS Software

Table S.2 shows the data conversion capabilities of some of the popular GIS software packages.

Table S.2 Summary of Data Conversion Ability in GIS Software

Software	File Format					
	TIGER	DIME	USGS DLG	ETAK	IMAGE	DXF
Arc/Info	yes	yes	yes	yes	yes	yes
MGE	yes	yes	yes	yes	yes	yes
GeoMedia	yes	yes	yes	yes	yes	yes
Atlas GIS	yes	yes	yes	yes	yes	yes
Geographix	yes	yes	yes	yes	yes	yes
TransCAD	yes	no	no	yes	yes	yes

CONCLUSIONS

In this report, typical transit GIS applications are reviewed. The data needs for various transit GIS applications are identified. The sources from which GIS data may be obtained are also identified and described. While developing transit GIS applications is not an easy task, and some basic training and understanding of GIS are necessary, it does not necessarily demand a full-fledged GIS department to implement some basic and simple applications. These may include demographic analysis of the service areas, analysis of route performance, maintaining records of paratransit users, customer information systems, etc. To implement transit GIS applications at a low cost, it is important to select a user-friendly GIS program, preferably desktop GIS, for the end users. Even for a large transit property where GIS expertise is available, desktop GIS will allow more transit professionals to easily access the GIS data and use the data in their daily work.

Another way to keep the cost down is to select the right data to use. The right data should have a locational accuracy adequate for the intended applications and the needed attribute information with good accuracy (e.g. correct street names, correct travel time). Before any field data are collected, existing data sources should be exhausted, such as the National Transit GIS Database and the TIGER/LINE files. It is likely that little or no field data collection needs to be done for implementing simple applications as mentioned before.

An important data issue is the maintenance of the data. First, there is always a lack of resources to constantly update GIS data or correct the data. Data may be rather outdated and still have to be used for current applications. Sometimes it may happen that one GIS layer is more current than another and the two do not agree with each other. For instance, a bus route layer may be current, but the bus stop layer may be several years old, resulting in some of the bus stops missing or some others not being served by any routes. The users should always be familiar with the data, including the source, data collection method, date of creation, accuracy, projection and coordinate system used, update history, etc.

Second, GIS data maintenance practice is not well established. There is often a lack of formal procedure for data maintenance. Data are sometimes maintained by individuals who develop them and/or use them. Due to change of personnel, the data history may be lost. It is therefore important to document all changes made to the GIS data, together with the original information describing the data (often called meta data), including both locational data and attribute data. Such documentation allows past changes to be traced and provides useful information to both database administrators and users.

To give transit properties easy access to information related to transit GIS applications, data sources, and software, the establishment of a central information center may prove effective. This information center can distribute information through the Internet and provide hot links to various Internet sites that contain relevant information. In addition, this center can provide basic training to transit professionals so that they have a basic understanding of the usefulness of GIS.

1. INTRODUCTION AND PROBLEM STATEMENT

Geographic Information Systems (GIS) have been recognized as invaluable information management and decision making tools for public transit planning and management. Many transit properties have made various levels of effort to use GIS in transit planning and operations. These efforts, however, are sometimes motivated by immediate needs and are often limited in their sophistication. There are many challenges to be faced if GIS is to be used widely and effectively in transit. Some of the major challenges, as identified in (Schweiger 1992), include funding, resources and training, lack of recognition of GIS capabilities or the value of GIS, and data issues.

Data are a critical element in GIS. GIS is data intensive, not only in terms of the tremendous volumes of data used, but also in terms of resources required for data collection and conversion, which often amount up to 80 percent of the total GIS investment costs. Due to limited resources and ability in the past to effectively utilize comprehensive transit data, many data important to transit planning and operations are often either not available, or, if collected, inadequate for GIS applications. In recent years, some of these data have become available or may become available with the introduction of advanced technologies such as automatic vehicle location (AVL) systems, automatic passenger counting (APC) systems, and smart fare cards. There is, however, limited understanding regarding what such technologies will offer in terms of a greatly expanded database enabling much more sophisticated and effective transit planning and operations analysis. In many cases, even when data useful for transit applications are available, due to a lack of trained GIS specialists available to transit properties, they are not utilized.

In the past, transit data maintenance is often done manually. In many cases, historical data are not well maintained. This presents a great loss of a rich knowledge source and opportunities to improve transit planning and management practices. For instance, trend analysis is not possible without historical data. While system level or route level ridership data are usually maintained and trend curves may be plotted, such trend information is overly simplistic since it does not reveal accurately the causes underlying the trends. GIS offers the opportunity to better maintain the transit data that may be used for planning and operations, as well as trend analysis. However, maintenance must also ensure data quality, currency, and integrity. Careful considerations must be given to the possible effects of data change in a GIS database on the various analysis functions that GIS supports.

Successful implementation of GIS for transit requires a good understanding of the aforementioned issues. While there have been studies in the past on transit information management systems (TRB 1994, Hill 1995) and the state-of-the-art transit GIS applications (Schweiger 1992, Moyer 1995, NUTI 1995), GIS data issues have not been adequately addressed. Existing literature on GIS standards (e.g. Racca 1995, FTA 1996) is more data- or database-oriented than application-oriented, addressing issues mainly on file formats, data exchange formats, and database design. Transit properties, especially small properties without the resources to hire GIS professionals, however, are in need of information on GIS applications and existing data sources to be able to implement GIS applications. Providing such information is therefore the goal of this research.

2. RESEARCH OBJECTIVES AND METHODOLOGY

This research is intended to identify the data needs for various transit GIS applications and possible sources from which useful GIS data may be obtained. The objective is to help transit properties, especially small and medium transit properties, which often lack the technical expertise and resources required by sophisticated GIS applications, better plan GIS implementations by providing information on possible GIS transit applications, the data needs for these applications, the existing and potential data sources, and other issues related to data quality and maintenance. This report, however, may also be used by transit professionals and GIS analysts as a reference to transit GIS applications and GIS data sources.

The research involved a comprehensive literature review. The purpose of this review is to identify the current GIS applications in transit, the data used, and the methods by which the data are collected. GIS standards are also reviewed.

To systematically identify the data needs for transit GIS applications, known and potential GIS transit applications related to transit planning, operations, and management are studied. Based on the specific applications or types of applications, the data needs and possible data sources are investigated, including those that may become available with the implementation of an advanced transit technology such as AVL and APC. The general characteristics of different data sources including the contents, accuracy, frequency of change, currency, maintenance, etc., are studied.

In the remainder of this report, **Section 3** provides a review of the literature on transit GIS applications. **Section 4** identifies the data requirements for various transit GIS applications. **Section 5** provides an inventory of transit GIS data sources. **Section 6** concludes the report with discussions regarding the needs for improving GIS data availability and GIS data maintenance practices.

3. LITERATURE REVIEW

GIS applications have been implemented at many transit agencies over the past 10 years. A 1992 study by Schweiger provides an overview of these GIS applications in public transit nationwide. Schweiger conducted 74 interviews with 67 organizations across 30 states, including 46 transit agencies, and 21 Metropolitan Planning Organizations (MPOs). The results of this study showed that GIS was being used or being implemented at that time for a wide variety of applications. Among the 46 transit agencies, 21 have GIS, and of the 21 MPOs, 15 have GIS. In 1995, another survey was conducted by Harman (1995). In the study, 269 entities were contacted with 202 survey forms completed. Later, all transit agencies and MPOs in urban areas with a population of over 200,000 were interviewed. Nearly every transit agency (TA) with a fleet of 500 and above indicated that they were using GIS in some form or another. GIS is used in transit agencies mostly in areas related to transit planning, transit operations, such as scheduling and run cutting, marketing, fixed route and paratransit dispatching, customer services, and management. MPO GIS uses are mainly for forecasting ridership, service planning, and development of map products.

In this section, various current and potential GIS applications in public transit management are described. These applications, reported in publications, represent the current state-of-the-art, but certainly not the limit of GIS technology.

3.1 Long Term Transit Planning

Azar and Ferreira (1995) examined the benefits of using GIS tools for developing transit ridership estimation models capable of forecasting changes in ridership, which may be associated with changes in bus route alignment and other modifications to transit service characteristics. The main interest in their study was to obtain a model that could predict transit ridership along a route based on the socio-economic attributes of an area, the physical characteristics of a bus route, and the attractiveness of "down-route" trip destinations using GIS. The Period Route Segment (PRS) Models, developed by Batchelder *et al.* (1983), was used for this purpose. This model estimates the A.M. peak and the midday boarding in each direction for every segment of a route based on route characteristics and service provided. According to the PRS model, ridership on a segment is a function of three factors:

1. A production factor related to the ability of an area to produce transit trips;

2. An opportunity factor representing the ability of areas down route to motivate persons to take these trips along that route; and

3. A level of service factor related to the quality of service provided along a route.

The PRS model requires computation of several variables that are spatially related. These variables, such as employment opportunities or people living within a quarter mile of a bus line or the stops that belong to the routes that generate transfer trips, may be easily calculated or approximated from maps with a GIS. A base map and 1990 census data for Boston, MA, were used to test the transit ridership estimation model. The 1990 TIGER/Line™ files and census block group boundaries were loaded into ARC/INFO® software along with population and employment location data from the 1990 census, selected ridership data, route alignments, and schedules for Metropolitan Boston Transit Authority bus routes. Using ARC/INFO GIS functions, the model displays the selected route, divides it into modeling segments, and shows the estimated A.M. peak and midday ridership on each modeling segment. Furthermore, the model estimates the total ridership on the whole route for both time periods.

Allen and Mukundan (1993) developed a process that used GIS to analyze different transit alternatives for major transit capital investment studies. Such alternatives included new fixed guideway facility, significant improvements in bus service frequency, new feeder bus routes, and new park-and-ride lots. The impacts of such alternatives were then evaluated.

Using GIS and artificial neural network (ANN) technology, the impact of demographics on fixed guideway ridership was studied by Shen *et al.* (1996). Demographics, fixed guideway alignment and stations, and ridership were correlated with both traditional statistical methods and ANN methods. ANN methods gave better results with smaller errors when compared to regression models. It was also an attempt to examine the trend of the demographic changes and its impact on fixed guideway ridership, but the effort was abandoned due to the lack of historical data in a GIS format. A similar effort of analyzing transit demand based on demographics was reported in (Wensley 1995). GIS was used to compute demographic characteristics for varying sizes of buffer zones around transit routes and stops, display the Census Transportation Planning Package (CTPP) data of origins and destinations of work trips, and generate transportation network for demand modeling purposes. Analyses were performed for a number of transit properties including New York City, Cleveland, San Juan, and Boston.

The interaction between land uses and transportation and transit systems was studied in (Eberlein and Brown 1991). In this study, transit and land use data were analyzed at the regional level using a GIS by relating the rail transit use patterns to siting of federal facilities in the Washington, D.C. area. From the study, recommendations were made to concentrate federal facilities or relocate some of the existing ones in order to not impact negatively the transit use and traffic congestion. This tool demonstrated that policies may be used as a tool to design land uses more rationally and promote transit use.

A study was done by Hsiao and Sterling (1992) to examine how geographic information system techniques were applied to analyze commuter rail survey data from which the upcoming intercounty rail service design may be projected and tailored. The application in this study demonstrated that GIS provided an efficient approach for transportation data analysis, particularly in the area of origin-destination data analysis. This study focuses on proper parking supply and traffic flow at rail stations, including the riders' origins and their boarding stations tabulated in O-D tables. These O-D tables were used as input to a network assignment for displaying traffic impact on the street network at rail stations. As a result, O-D trip tables were applied to the travel forecasting model for network assignment to the peak-period link volume. It became possible to gauge the distribution pattern of a rider's origin in relation to boarding stations. By analyzing the network assignment, it was found that most people traveled three to seven miles from their homes to the rail stations. The authors concluded that the GIS provided a flexible approach for data analysis for the specific purpose. Moreover, both short-term commuting service design and long-term range transportation planning activities may be developed based upon a common GIS database structure, from which specific information can be queried to address a unique project purpose.

Broward County, Florida, conducted a demonstration project called the Broward County Transit Options Project (TOPS). Funded by the Federal Transit Administration, the project was aimed at developing a variety of travel options using public transit including paratransit, fixed route, local community bus services, and accessible private taxi services. The latest GIS technology and data resources were used to investigate the usefulness of the state-of-the-art technologies in the design process. The approach used to apply GIS to integrated service delivery, strategies for data acquisition, and strategies for data analysis are presented in the final project report (Lawrence and Russell 1995).

Another application of the use of GIS in transit can be the development of an intermodal transportation plan. The Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 requires that each state prepare a statewide intermodal transportation plan in mid 1995. Louisiana, along with Alaska, Florida, New Mexico, Ohio, and a consortium of six New England states were selected to develop modal plans. In this study, Movassaghi and Parlee (1995), Louisiana Department of Transportation and Development took into consideration not only the passenger movements, but also the cargo movement including road, rail, waterway, and air travel. By determining the type, level of detail, and the methodology to perform the analysis, data were collected. The data collected included network and flow data (base-line 1990 and forecasts 2020). Network data includes demographics, land use, highway network, rail network, air network, and waterway network, while flow data includes node network for passenger and cargo movements.

Other non-spatial data are stored in MIS separately. Such data includes ozone and carbon monoxide concentration in air, major imports and exports of various regions, emissions by vehicle type, fuel consumption by vehicle type, operating cost by mode, freight shipment rates by commodity, regulation and legislative acts, results of interviews with transportation users and providers, and average vehicle occupancy by mode. The study concentrated primarily on the intercity movements. The database development evolved into analysis and mapping results for future flows of passenger and cargo. These results were then displayed on maps, and based on which additional highways were proposed.

In 1989, the Metropolitan Transportation Authority (MTA) in New York City used TransCAD to support data management, modeling, and decision support for transportation planning. The data collected during this project were used for the highway and subway networks with census tract polygons containing information on population, income, and other socioeconomic characteristics (Prastacos 1992).

3.2 Service Planning

An application of GIS to bus routing is described by Chou (1995) regarding the design and application of a decision support system developed for bus routing, route sequence mapping, and passenger geocoding. This application, although specific to school bus routes, may be extended to include revenue service bus routing for both fixed route and demand response services. The school bus routing system consists of six operational modules:

1. A Single Routing Module allows planners to select either street segments or bus stops from a street map to generate his or her optimal route. The routing information includes the street address associated with each segment, the distance and estimated travel time of the segment at the assigned speed limit, and a turn code indicating if a turn is required at an intersection. A route may also be determined by selecting a sequence of bus stops.

2. A Walking Distance Maintenance Module identifies street segments that are within a user specified distance from a school. This model helps the school district to identify the students who live within a specific distance from schools and thus are not eligible for the school bus service.

3. A Bus Stop Optimization Module identifies the optimal sites for the location of bus stops based on travel demand.

4. A Passenger Plotting Module reads the passenger information file, geocodes it, and then plots their locations on the street map. The function of this module is to display the geographic

distribution of any selected group of students in the school district. It also allows the user to retrieve the student file, specify any combination of selection criteria, such as a student of a certain school or within a specific range of grades, then plot on the street map the residential location of all students that satisfy the selection conditions.

5. A Multiple Routing Module defines multiple routes based on information of passengers, bus stop location, and the bus fleet. The module assigns each available bus to serve a sequence of closely located stops until either the bus is full or the travel distance exceeds a limit.

6. A Complex Routing Module generates multiple routes with multiple destinations (schools) while each bus may be assigned to more than one route. Some constraints are incorporated into this module such as the availability of handicap equipment, the maximum acceptable time a passenger may stay on board a bus, and the bell time of each school. This module also allows students to go to one school on certain days and another school on other days. The conceptual framework of the bus routing system is shown in **Figure 3.1** Additional modules may be programmed to serve specific users. The routing system is designed to generate optimal solutions based on the existing locations of bus stops and the distribution of travel demand. For planning new bus services, the system optimal locations of bus stops may be identified according to the geographic distribution of the passengers. Other applications may require some user-specific or system-related constraints such as the seat capacity of each bus and the maximum loading factor for each segment. The data for this system requires a digital map of the street network, the location of schools and bus stops, a file of bus fleet, and a file of passenger information including each passenger's street address. The complex routing module requires additional information about any disabilities of each student and the equipment available on each bus assigned to the special education program.

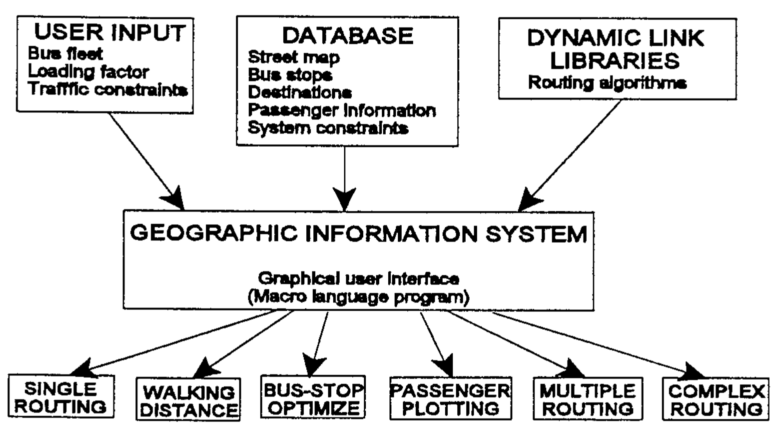

Figure 3.1 Conceptual Framework of Bus Routing System (Chou, 1995)

The use of GIS in planning transit services for people with disabilities was discussed by Javid and Prabhakar (1993). This application is much simpler than that of the school bus. The difference is to develop a method to estimate the number of people with disabilities within the catchment area (acceptable walking distance) or transit service area for the purpose of estimating passenger demand. The information for building the database was obtained from TIGER/Line™ files, and from a questionnaire survey conducted as part of another study. Data from TIGER/Line™ files include county, state, census tracts, census blocks, blocks groups, intersections, and roads. These data include total population and disabled population by block group. The origin destination trips of the people with disabilities surveyed from a previous study were entered into a point database. To represent the transit routes, separate databases were built for each direction of the six fixed routes operated by the Logan Transit District (LTD), Cache County, Utah. The point database and the routes database were then used to estimate the transit demand for each route and the entire network.

The GIS database may also be useful to schedule transit service to accommodate a planned event (Peng et al. 1995), such as a special event or a large festival. Once the location of the special event is identified and geocoded, the GIS will be able to schedule special bus services to link the event location with the fixed route services as well as to make changes on the level of service on the original fixed route services.

In a report completed by the Community Transportation Association of America (CTAA 1998), the use of GIS to analyze transit needs of welfare recipients is described. Joseph Coughlin, director of the New England University Transportation Center at the Massachusetts Institute of Technology, and Michael Rich, associate professor of political science at Emory university in Atlanta, plotted locations of entry-level jobs advertised in newspapers, residences of welfare recipients, sites for child care and job training facilities, and availability of transportation in Atalanta metropolitan area. They found that only 43 percent of entry-level jobs in Cobb county are accessible to welfare recipients by public transit. In addition, most of these jobs entail one- to two-hour commute time.

The same report also describes an effort by the Department of Social Services of St. Mary County, Maryland, to identify welfare recipients' residences and employment opportunities. The public transit services are then analyzed to identify the need for new services or service improvements.

3.3 Transit Operations

In transit operations, vehicle dispatching, routing, and fleet management are a few of the activities in which GIS may offer help. For fixed route services, vehicle dispatching and routing are mainly for vehicle maintenance (e.g., returning a malfunctioning bus to maintenance facility), deadheading, and emergency responses. For paratransit, efficient dispatching and routing directly impact the operating cost, service effectiveness, and customer satisfaction. In (Yang 1995), technical approaches to and issues related to dynamic vehicle routing problems (DVRP) are discussed. A prototype system is also introduced. According to (Yang 1995), using GIS for paratransit routing and scheduling, vehicle miles may be reduced by 20 to 30 percent. GIS applications for transit operations including display of passenger boarding and other business performance on maps is discussed in (Reilly 1997).

In recent years, automatic vehicle location (AVL) systems have been installed on transit vehicles in many transit agencies, which, by providing vehicle location information in real-time, allows the dispatchers to monitor the fleet continually, warn the drivers about delay in schedule, and respond immediately to emergencies. Most AVL systems have a component that is a GIS-based display of vehicle location. Queries may be made regarding the exact location or address of a transit vehicle and its status. In (Raj et al. 1995), a GIS-based emergence response system is described that provides the dispatchers with accident/incident dependent, sequential and prioritized action plans when an accident or incident occurs. It also records the actions taken by the dispatchers and files incident reports.

3.4 Transit Safety and Security

Transit safety is of most importance to transit users. Due to traffic congestions, movement conflicts at, e.g., railroad crossings, and other potential disasters such as hazardous material spills, risks of accidents and incidents must be minimized. Panchanathan and Faghri (1995) discussed the development of a knowledge-based GIS for improving rail-highway crossing safety. The GIS interfaces with the USDOT model and a knowledge-based expert system, and, making use of qualitative and historical information, evaluates the safety at a given crossing. Many highway safety GIS applications have been developed that may be used for transit applications with minor modifications.

The use of GIS for road safety analysis and management in Haifa, Israel, was presented by Peled and Hakkert (1993). Although the GIS application was not specifically developed for transit, because transit modes that share rights-of-way with general traffic experience the same safety problems as automobiles, the application may be extended to meet transit needs. Road safety analysis may be divided into three different processes:

- Identification of a problem;
- Diagnoses of preventative measures; and
- Selection of a solution.

Analysis of accident data may result in several possible actions:

1. Modification of a roadway facility by raising engineering standards on, for example, lighting, friction, guardrail, etc.;

2. Single site and route actions associated with the identification of safety problems on a specific route or a segment of a route; and

3. Area action that deals with problems associated with the geographic area such as a neighborhood, a region, a district, or a municipality.

The data for this study was collected from different sources. A data file compiled by the Planning Department of the Israel Police was suitable for work only at the municipal level because it has no location code. A data file was prepared by the Central Bureau of Statistics that compiled accident data by urban area, rural area, and towns. In this file, accidents for each town are associated with a road section or intersection. A local accident file specific for Haifa was fed directly from information supplied by the local police department. Difficulties facing the police department in collecting accident data were inaccuracies in data, lack of information, and incomplete or untimely reporting of accidents.

Accident data alone was insufficient for the analysis. Accident data needed to be analyzed to consider the time, location, and the geometry of the road where the accident occurred. In this context, GIS was the solution for better accident and road safety analysis. The two major databases involved in the road safety analysis are the road system layout and accident information. A few other databases are also used, including intersections positional and physical characteristics, traffic counts, intersection related data such as signals and signal programs, traffic signage, pavement markings, and public transportation services and routes.

Using GIS software (ARC/INFO), two layers of road system positional data (streets and intersections) and one layer of accidents were created. The accident layer is complemented by two additional non-georeferenced (tabular) data files of vehicle and casualties involved in the accident that are related to and controlled by the accident layer. A special landmark layer was also defined to elaborate on analysis pertaining to attractions like schools, cinemas, supermarkets, etc. Photographs for several intersections and links were scanned and added to the system in order to develop a tool allowing users to view places of interest such as hazardous intersections, bus routes, etc.

This GIS is capable of generating reports in a tabular format or map format. These reports enable managers to understand the overall accident situation in their area or at a particular location, and how it is changing at frequent intervals. Furthermore, they allow the staff members to determine traffic operational and capital investment solutions. These reports establish priorities for improvements on a regular basis. The geographic accident database described above is structured in such a way that the total number of accidents and accident rates at an intersection, on a street segment, or an entire street length, or in a specific area may be determined. The information obtained from the periodic reports and map displays, combined with an accessible database can assist staff members to respond efficiently to citizens' complaints.

In 1980, the Ohio DOT (ODOT) began to develop a GIS (Gebhardt 1992) to initially display accident reports to the nearest one hundredth of a mile on maps of state, county, township, and municipal jurisdictions. All jurisdictions were digitized with total road millage of approximately 112,000 miles. Another task was to relate all digitized roadway mileage to the truly traversed roadway mileage used in the Road Inventory. By adding inventory data to the database, ODOT were able to perform analysis for the desired information which could be displayed graphically with plots or interactively on screen.

The Kansas Department of Transportation (KDOT) developed a GIS accident analysis system using a CAD package and accident data files (Schweiger 1992). The city of Boulder, Colorado, was in the process of implementing GIS for transportation management (Schweiger 1992). The end product would allow access to traffic counts. Furthermore, it would provide needed information on accident, traffic control device, traffic safety problems, and maintenance needs.

Berman *et al.* (1997) described the development of a GIS desk top application for transit security incident analysis and reporting in the King County Department of Transportation's Transit Division. More than 3,000 security incidents per year have been tracked and reported in the past using a DOS-based database management system. The objectives of developing a GIS system were "to maintain security for drivers and passengers, identify systematic patterns in the occurrence of security incidents in order to optimize the deployment of security assets, provide timely information on security-related incidents, and provide historical information in support of planning special event coverage." Implemented using MapObjects and Microsoft Visual Basic and incorporating GIS technology with standard database querying capabilities using Microsoft Access, the system supports incident data entry, analysis, and reporting. A variety of spatial and nonspatial database queries may be made through a graphical user interface.

Samples of GIS applications for crime analysis, prevention, and visualization may be obtained from (Agung 1997) and (Chou 1997).

3.5 Service Monitoring

A study was conducted by Papacosta (1995) for the Honolulu bus system, Oahu Transit Services (OTS), to monitor the schedule adherence for the fixed route bus systems operations using GIS. The data incorporated in this study include ride checks and point check surveys collected as part of a comprehensive operating analyses. Each bus trip sampled was computerized in a particular format and placed in a separate ASCII file. Each bus trip file includes information such as the date of survey, checker name, route number, direction, key and run, bus ID number, and seating capacity. A list of the bus stops with the corresponding arrival time and boarding/alighting is also included. A program written in C language converts the data into the Arc/Info file format. Program modules for this project were written using Arc/Info Macro Language (AML). The analysis program obtains and processes the necessary data for the required run. For example, when the user selects a route key, the program identifies the route structure and its corresponding stops. It then searches the database for all scheduled runs associated with that key, obtains the necessary raw data, and stores them in a temporary file for further processing. Any requested displays appear in separate windows on the output device. Hence, they can be viewed and examined jointly.

3.6 Customer Information Services

GIS may be used as a powerful tool to provide the transit users with information about transit services on a digital map that may be used for pre-trip, in-terminal, or in-vehicle planning. Using a modem connection, travelers may connect their personal computers with the transit agency's telephone information center and display a transit map. Another way to dispense travel information is through kiosks at public places such as universities, government offices, and shopping malls for planning non-home based trips. By specifying the trip origin and destination, information such as route(s) to be used, arrival time at each stop, transfer points, fares, nearest parking lots, ADA accessibility, etc., may be obtained. Automatic trip planning systems have been installed in many transit agencies. Among transit agencies with such a system are: Winston-Salem Transit Agency in North Carolina, Los Angeles County Metropolitan Transportation Authority (LACMTA) in California, Tri-Met in Portland, Oregon, Metro-Dade Transit Agency (MDTA) in Miami, Florida, Metropolitan Atlanta Rapid Transit Authority (MARTA) in Atlanta, Go regia, and Central Ohio Transit Authority (COTA) in Ohio. Recently, Tri-County Commuter Rail Authority in South Florida recently begun to implement a real-time travel information system that employs GPS to track trains and that will display train schedule on message baords at stations.

In-terminal information systems provide schedule updates and transfer information for passengers. This information includes arrival and departure times, information on transfers and connections, information on other regional transportation services, and park and ride facilities. On board of a vehicle, information on routes, schedules, and connecting services are provided via visual displays and annunciations. Transit routes and schedules are needed to provide such information. For more accurate information, an automatic vehicle location (AVL) system may be used with GIS to provide planners, operators, and users with real-time information. Scheduled time and route information may be adjusted according to traffic delays, accidents or emergency situations. In case of an emergency, a vehicle may be easily located and emergency vehicles timely dispatched.

Tri-Met installed kiosks for trip planning, mainly for visitors in the downtown area. The kiosks contain microcomputers and allow users to enter their destinations. They will then identify routes and provide directions. A display of bus routes serving the area in the vicinity of the user's origin and destination will be displayed.

3.7 Marketing

Henry and Sirota (1995) described an effort of identifying ways of increasing transit ridership on the Baltimore light rail system. Selected management level employees were asked to identify the areas with market potentials and develop recommendations to attract transit users. GIS was used to display and analyze ridership, origin-destination data, and CTPP data. Lycan and Orrell (1990) also described a GIS-based approach to analysis of the potential of bus ridership to Oregon Sciences University. Address matching was used to analyzed the accessibility of the university to the commuters. The results helped to understand the feasibility to modify or improve the transit services to the commuters.

3.8 Facility and Real Estate Management

Facilities directly owned or managed by a transit agency typically include right-of-ways for fixed guideway systems, wayside equipment, storage yards, administrative and maintenance buildings, transit stops, stations, signs, parking lots, land parcels adjacent to transit stations or transit centers, and special structures dedicated for transit use including bridges and tunnels. Other infrastructure that is used by transit but not owned or managed include roadways, expressways, highway bridges, drainage systems, traffic signals, general traffic signs, seaports, airport, interurban train stations, railroad tracks, etc. In (Cipolloni 1993), the use of GIS for improving the management of transportation facilities were discussed. An example of how New Jersey employed this technology to manage their traffic signs is also given in some detail.

Sathisan *et al.* (1993) demonstrated that GIS may be used for the management of rail infrastructure. This study addresses the development of a GIS-based system for rail infrastructure with respect to the key element and characteristics of a typical mainline railroad track, such as:

- Track alignment;
- Longitudinal gradient and elevation;
- Sidings;

- Bridges and culverts; and
- At-grade rail/highway crossing.

The paper also discusses the application of such a system for risk and routing analysis as related to hazardous material shipping. A GIS coverage of the U.S. Census Bureaus's TIGER/Line file rail alignment in Clark County, Nevada, was used to serve as a starting point and to provide a base map. The next step was to relate all infrastructure characteristics to a specific milepost location. The accuracy of these mile posts was one-hundredth of a mile. Other data required for the development of the database were obtained from Union Pacific's Western District Condensed Track Profile (1985) and Western Area Condensed Track Profile (1990). After the development of the database and the graphical display of the rail infrastructure, probabilistic and quantitative risk assessments of alternatives were conducted to minimize the risk of routes transporting hazardous materials.

Ryu *et al.* (1997) describes a GIS application for management of urban transit facility for the Pusan Urban Transit Authority (PUTA), which operates an extensive network of underground transportation systems since 1987. To effectively manage the infrastructure, about 40,000 sheets of existing A0/A3 size drawings were scanned and vectorized, from which GIS layers were constructed. The system stores information about facilities such as railway, tunnel, electricity, signal, communications, and others such as administrative boundaries, geographic features, drainage, and raster maps.

3.9 GIS in Transit Agencies

Other opportunities exist for the use of GIS in transit planning, operation, and analysis. Some transportation organizations have become extremely active in the use of GIS. Smaller organizations have been less likely to invest the resource necessary to establish a GIS that will generate significant benefits (Harry 1995). The following discussion represents a selection of transit systems that use GIS in transit planning, operation, and/or analysis.

King County Metro Transit Municipality of Metropolitan Seattle

King County Metro Transit in Seattle, Washington, is one of the most active users of GIS among transit agencies in the United States. In 1982, King County Metro developed an in-house GIS called TransGeo. As of July 1992, TransGeo continued to be a critical feeder system to Metro's automatic passenger counts system, ARIS/BUS TIME, commuter information system, mileage maintenance system, and the radio automatic vehicle monitoring/location system. In 1991, a GIS project team was established to carry out a GIS project that involves the assessment of the GIS needs of the agency and analyses of alternative GIS implementation strategies. As of today, GIS applications have included (CUTR 1996):

- Capital planning and development: display park-and-ride data against service patterns, volumes, and passenger volumes.

- Service planning: produce maps of selected service areas, and display route(s) proposals and information related to ridership demand such as population, employment, travel patterns, population, and employment densities.

- Market development: display employer sites against available transportation services and information such as transit service, park-and-ride, etc.; display and analyze demographic, census data, etc.

- Accessible services: geocode ADA applicant addresses and perform location related analysis.

- Coach and facility maintenance: provide route pattern maintenance (to track mileage by vehicle and route).

- Sales and customer services: use geocoding to create customer mailing lists for route-level marketing.

- Research and market strategy: create displays and other test materials for focus groups; geocode respondents' origin and destination to create a travel pattern database.

- Transit operations: create quick information maps/guides for operators to help answer customers' questions; provide timely and accurate maps for trainers and operators.

- Risk management, safety, and security: display and analyze accident locations by various attributes.

- Communications and community relations: use GIS to convey information to decision-makers and the general public.

- Rational transit project: calculate coverage of services against population.

- Environmental planning: display environment elements such as transit routes, sewer/storm systems, rivers, lakes, bays, housing patterns, etc.

- Analysis of accidents by various attributes.

Dallas Area Rapid Transit (DART)

DART's GIS is a critical planning and management tool for anticipating network service needs, and therefore, for improving existing and future ridership. GIS applications at DART are used in numerous divisions of the agency, including the following (Schweiger 1992, CUTR 1996):

- Paratransit;
- Automatic vehicle location;
- Enhanced customer services;
- Bus stops inventory management;
- Service planning;
- Data collection
- Service scheduling; and
- System-wide analysis.

The purpose of developing and implementing a GIS at DART was to respond more quickly to customer requests, effectively track route changes, and to determine required additions and modifications of transit lines. DART's GIS is used primarily for service planning, customer assistance, and map and schedule production. For instance, GIS was used extensively to perform a series of studies to examine three different rail alignments in north Dallas. Current and projected population and employment densities were overlaid and buffer analysis was performed.

In March 1993, Dallas Area Rapid Transit (DART) began implementing TIPS, which allows information operators to access a trip's origin or destination by inputting an address, street intersection, or landmark. The system geocodes the location and searches for the closest bus stop to that location, and reports all routes that stop at the location and displays the scheduled information. It can be developed to calculate the customer's entire trip and provide alternatives.

Los Angeles County Metropolitan Transit Authority

LACMTA provides transit services to Southern California residents and visitors. The MTA maintains an integrated transportation network called the Metro System, which includes bus and rail services, transportation demand management, bikeways, and highway improvements. LACMTA has been implementing an agencywide enterprise GIS to better manage the vast array of GIS activities, including managing existing transportation networks, long range regional planning, emergency preparedness, and property management (Mumbleau and Davis 1997). LACMTA also planned to use GIS to predict the future road congestion and the projected patronage on a proposed subway. Other areas of GIS applications will be providing solution to traffic problems, customer information systems, tracking and assessing industrial and commercial properties to support the Metro Rail Program (Peng *et al.* 1995).

For long term planning, TRANPLAN is used for transit modeling and Arc/Info is used for GIS analysis of modeled forecast results. The two packages are linked with data conversion software provided with TRANPLAN. Arc/Info's spatial analysis capabilities are used to merge data sources with modeling scenarios in a graphic environment. Transportation zone models are developed with Arc/Info and then transferred to TRANPLAN for analysis. The resulting models are fed back into Arc/Info, displayed, thematically mapped, and plotted. By linking modeled output to the GIS, results such as future road congestion and projected patronage on a proposed subway are easily visualized.

San Diego Associates of Governments (SANDAG)

SANDAG has been using GIS since 1985, and has developed an extensive collection of spatial database (CUTR 1996). Transit planning and marketing are key functions of SANDAG's GIS. Databases of population, housing, employment estimate, crime statistics, and land use are maintained and used to produce historical, current, and forecasted profiles. The use of GIS in transit planning allows sophisticated analysis of complex transit-related queries using large databases. The use of TRANPLAN with GIS provides a powerful tool for transit modeling. Recent work includes the development of transit flow maps which provide users with a simplified visual representation of ridership volumes by route. Decisions regarding where to expand or reduce service may be assisted by these maps and plots. Table 3.1 represents the different GIS applications at SANDAG.

Milwaukee County Transit System

Milwaukee County Transit System (MCTS) is operated by Milwaukee Transit Services, Inc. The GIS of Milwaukee County Transit is integrated with computer-aided dispatch (CAD) and Automatic Vehicle Location (AVL) bus system (CUTR 1996). The Westinghouse system (SmartTrack) provides reliable two-way radio communication. Information from SmartTrack helps drivers and dispatchers proactively manage fleet operations and maintain arrival and departure schedules. Also, daily activities such as bus operator requests for vehicle repair and security assistance are tracked. Time point and schedule adherence information is also logged into a database so that the scheduling department can monitor trends in on-time performance.

Future applications for MCTA include the following:

- Installation of automatic passenger counters are planned for fifty buses by 1996;

- Options for the purchase of mapping/census software for the Planning Division was to be investigated; and

- Schedule adherence information will be available to planners for review of schedule running time.

Table 3.2 summarizes transit agencies using GIS for transit applications.

3.10 Other Transportation Related GIS Applications

Examples of using GIS in transportation corridor planning include studies in Dallas, Northern Virginia, suburban Atlanta, and Logan, Utah. Recent applications in North Carolina, within the Charlotte area, involve development of noise contours along roads to parcels that may be suitable for industrial development as opposed to residential development. Tennessee also developed a GIS virtual reality to view how newly proposed roads would fit into the landscape. This type of application may be extended to study noises near proposed or exiting fixed guideway systems including commuter rail, heavy rail (especially if at-grade or elevated), light rail, and busway systems.

Hartgen and Li (1993) presented the use of GIS in transportation corridor planning. North Carolina uses GIS to identify the major impacts of I-40, a 120 mile highway connecting Wilmington to Raleigh. GIS was used to provide recommended actions and strategies for governmental as well as transit agencies. It was also used to reduce negative impacts and maximize positive impacts. Extensive information was used in this study including:

- Demographic and socioeconomic data for all counties along the route;
- Traffic conditions;
- Population and household statistics;
- Information describing the business responses to telephone surveys; and
- Citizen information obtained from public hearings and forums.

This information was used with GIS to analyze the extent to which each of the 22 exits on I-40 were accessible from surrounding communities and other traffic corridors and which communities within the corridor would be affected negatively by greater accessibility to big cities. During this project GIS-T served as a basic tool for facilitating the communication between planners of data gathering, citizens, and policy makers.

Table 3.1 Base Applications for Desktop GIS at SANDAG

Application	Description	Category	Database	Status
Transit Service Potential	Using a definition of potential transit ridership including employment, income, low auto availability, renter, age, and other variables, define areas undeserved or unserved by transit	Planning/Operations	Census data	Currently done with SANDAG's GIS
Socioeconomic profiles of areas surrounding transit	Socio-economic profiles of areas surrounding stops, routes, route segments. Allows staff to buffer an area within a specified distance to a stop or route	Planning/Operations /Marketing	Census data Employment Inventory Transit Coverage	Currently done with SANDAG's GIS
Physical characteristics of transit	Maintains physical characteristics of bus/trolley stops, displays ADA accessible stops, etc. (i.e., What is the distribution of accessible bus stops?)	Planning/Operations	SANDAG and operator bus stop inventories	Currently done with SANDAG's GIS
Route analysis	Analyze existing and planned routes by stop activity, capacities, analysis of passenger loads by route segment	Planning/Operations	Transit coverage, route alternatives, Passenger counts, surveys	Currently done by individual operator
Title VI evaluation	Identify minority areas, transit accessibility, and minority routes for FTA/Title VI requirements	Planning	Census data, Passenger counts, Transit coverage	Currently done with SANDAG's GIS
Future Growth areas	Identify areas of forecast growth and relate these changes to transit (current and planned service). For example, what is the expected population growth within the service area of a planned light rail line?	Planning/Operations /Marketing	SANDAG population, employment, land use forecasts, Transit coverage	Currently done with SANDAG's GIS
Transit use by time	Identify route activity during peak versus non-peak hours, identify possible turnarounds	Planning/Operations	Passenger Counting Program, Transit Coverage	Not yet developed

Source: (Culp 1994)

The application of GIS-T in the Carolina Parkway study was an example of using GIS for forecasting transportation demand and network modeling. This project was part of a transportation study intended to coordinate land use and transportation planning in a way that creates an attractive, efficient regional transportation system supporting economic development objectives. The goal of this study was to develop a series of traffic and land use forecasts for the years 2010 and 2030, with and without the parkway. Several comparative analyses were prepared, showing traffic on critical road segments in the area. Tables for vehicle miles traveled, speed, vehicle hours traveled, and emissions were also developed.

Anderson (1992) presented an application of GIS to transportation planning for the Montgomery County in Maryland which is one of the early users of GIS technology for integrating transportation planning database. For transportation modeling development and data management, a GIS Spatial Analysis System (SPANS) was acquired. SPANS was used to produce maps of TAZ level accident of a certain level of severity in relation to pavement conditions.

input, to provide land use data input to a series of suburban planning modeling. It was also used in the analysis of the household travel survey especially the time-of-day travel behavior. The input data for Montgomery County includes 255,000 recorded county tax assessor parcel file, along with a file containing all approved subdivisions in the county, and vector graph files of parcel boundaries. Also included are TAZ boundaries, MetroRail and commuter rail stations, TIGER/Line street vector file, and sidewalk and street mileage file.

The Office of Policy Development (HPP-1) was the first office within the Federal Highway Administration (FHWA) to use a GIS for highway policy analysis (Stokes and Marucci 1995). HPP-1 started constructing its database in the mid-1980s and completed the database in 1988. The database coded 370,000 miles of highways. The attributes of the database include FHWA functional classification, route number, length, median type, access control type, number of lanes, and pavement type. Other databases in the GIS include interstate truck volumes, state boundaries, congressional district boundaries, airports, military installations and bridges.

Among the leading DOTs in the implementation of GIS besides Wisconsin and North Carolina was Pennsylvania DOT. It formed a steering committee consisting of several state agencies in charge of undertaking research to develop a strategic plan to implement GIS by the agency involved. GIS application in PennDOT is expected to follow the traditional areas of an agency responsible for pavement management, traffic engineering, safety, planning and programming, bridge rehabilitation, etc. (Basile et al. 1992). The first application for data retrieval was related to the highway safety program, in which accident records can be displayed graphically in many ways on the state highway system. The display of the accident was very helpful for police and emergency medical services. Data integration involved the use of two or more databases to develop the required information. For example, accident records may be combined with road data to view an

In a study for the Center Region of Pennsylvania, Matzzie and Rogers (1991) utilized MapInfo GIS in conjunction with MINUTP transportation modeling system. Transportation network modeling generally involves a large number of elements. Usually there are several hundred traffic analysis zones, and thousands of links and nodes, depending on the geographic extent of the area of study and the degree of detail incorporated into the model. MINUTP is a derivative of the UTPS software package that runs on mainframe computers. It includes two graphic interfaces that provide a powerful tool for viewing and editing transportation networks. Moreover, it provides a colored hard copy output of a network with different attributes for each link. Although MINUTP had provided the ability to visualize the network and help the user in model calibration, interpretation, and communication of the model, it did not show either the boundaries of the TAZ nor the network in relation to physical features. By using GIS as a database management tool and applying location codes to all records, information may be overlaid and a variety of location-oriented questions may be addressed.

Other technologies will become an important element of GIS such as the Global Positioning System (GPS). The use of GPS will enable transit operators or transportation professionals to quickly determine the location of interested features on the earth's surface that constitute the spatial data in a GIS database. GPS-derived data may be field recorded and inserted directly into a computer database without further processing. The Texas State Department of Highway and Public Transportation (SDHPT) has installed five automated GPS continuous tracking stations, located in Dallas, Austin, San Antonio, Corpus Christi, and Houston, to facilitate existing and future GPS operations. Among the non-highway transportation users, the railroads have already considered the possibility of using GPS to keep track of the railroad cars and to keep real-time accurate positions of their trains. The use of GPS with GIS is now being tested by the U.S. Coast Guard for harbor navigation. Moreover, FAA is evaluating GPS for both airborne enroute navigation and precise terminal approach applications.

Table 3.2 List of Transit GIS Application Areas

APPLICATION AREAS	USERS
Transit ridership forecasting, service planning, market analysis	Baltimore MTA, Bay Area MTC, Bi-State, Chicago RTA, City of Phoenix, DART, DVRPC, GGBHTD, HGAC, Houston Metro, LIRR, MARC, NJTA, NOACA, NYCTA, NYMTA, Omaha-Council Bluffs MPO, Port Authority NY&NJ, Portland Metro, SANDAG, Seattle Metro, SEMCOG, SMART, Tampa Urban Area MPO, TTD, WashCOG
Map products design and publishing (for example: system maps, route schedules and maps, operator maps)	Bi-State, City of Phoenix, CTPS, DART, DVRPC, Houston Metro, MARC, Miami MPO, MTDB, NJT, NOACA, NYCTA, Omaha-Council Bluffs MPO, Port Authority NY & NJ, Portland Metro, RVTD, SANDAG, Seattle Metro, Tampa Urban Area MPO, WashCOG
Fixed facilities and real estate management (for example: bus stops, transit stations, park and ride lots)	Balt.. MTA, Chicago RTA, City of Phoenix, DART, DVRPC, GGBHTD, Houston Metro, LIRR, Miami MPO, NYCTA, NYMTA, RVTD, SANDAG, Seattle Metro, SMART
Telephone-based customer information services	City of Phoenix, DART, DVRPC, NOACA, NYCTA, Seattle Metro
Transit scheduling and run time	DART, DVRPC, NJT, RVTD, Seattle Metro
Ride matching (for car and van pools)	DART, MARC, Miami MPO, Seattle Metro
Automatic vehicle location	DART, NJT, Seattle Metro
Transit pass sales	DVRPC, MTDB, Seattle Metro
Paratransit schedule and dispatching	TTD
Improved transfer points and connections	NYCTA

Source: (Schweiger 1992)

4. DATA REQUIREMENT FOR TRANSIT GIS APPLICATIONS

Data are a crucial element in GIS. GIS is data intensive, not only in terms of the tremendous volumes of data used, but also in terms of resources required for data collection and conversion, which often amount up to 80 percent of total GIS investment costs. In this section, the data needs for different transit GIS applications are described. The discussion is organized by major transit function areas, as shown in **Figure 4.1**. The transit functions, possible GIS applications, and data needed for such applications are described.

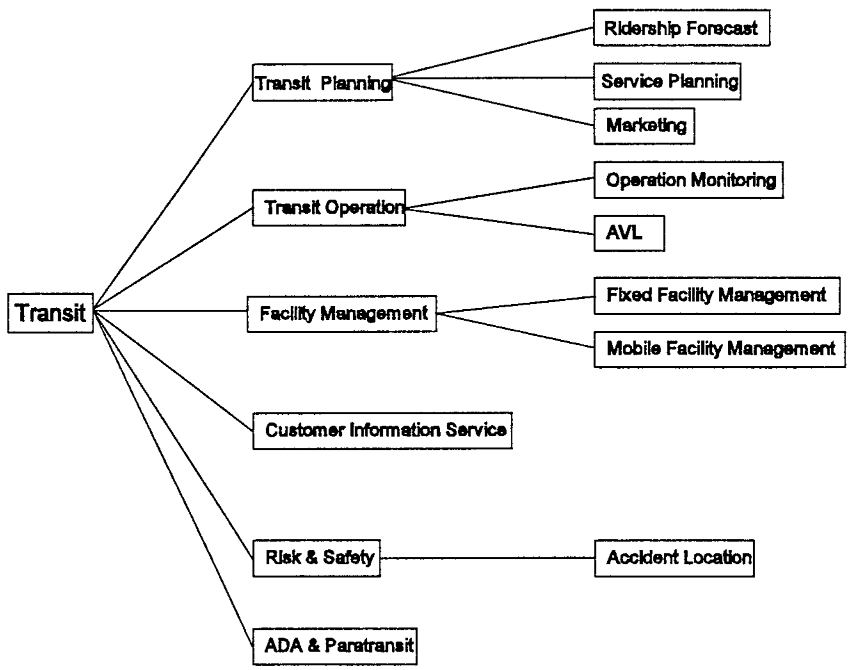

Figure 4.1 Transit GIS Applications

4.1 Transit Ridership Forecast

Long-term transit planning is needed for expanding existing transit systems or developing a new capital intensive facility (Levinson 1990). One of the most important reasons for changing routes or services is the change in ridership. Ridership influences route structure and service density, and is one of the primary basis on which a decision regarding a new major capital project is made. Accurate ridership forecasts are therefore of great importance.

A variety of models have been developed for ridership forecasting. Although these methods differ, most of them use data such as number of households, family income, and car ownership in the forecasting process. As described in (Levinson 1990), ridership estimation involves the following:

(1) Estimate the dwelling units and population within the coverage area, preferably by block. The data used for this purpose may be census data, postal area data, or data obtained from aerial photographs;

(2) Obtain the demographic and socioeconomic characteristics of the area. These include family income, type and age of dwelling units, and car ownership;

(3) Conduct surveys to identify market and travel attitudes and patterns;

(4) Obtain ridership potentials for residential areas served; and

(5) Identify schools, shopping centers, offices, and industries within the service area. Obtain estimates of enrollment, employment, places of residence, and probable travel modes. The ridership potentials brought by these activities are estimated and add to the residential ridership identified in Step (4).

GIS may be used in transit ridership forecast (or long term transit planning) in the following ways:

(1) Network building, data creation and extraction for transportation analysis zones, and spatial aggregation of data (Spear *et al.* 1994). GIS may be used to build transit network over the general transportation network and obtain demographic and socioeconomic data from sources such as Census Bureau. Demographic and socioeconomic data may be aggregated at census tract level, traffic analysis zone level, census block group level, or even census block level.

(2) Analysis of route level demographics, such as displaying population living within a certain distance of a particular transit route. Traditionally, the buffer zone method is used. It involves creating a buffer zone of a certain size around transit stations/stops or a route. The population (possibly of certain demographic or socioeconomic characteristics) or employment in the buffer zone area(s) is then estimated by proportioning the total population (assuming their even distribution) based on the ratio of the buffer zone area and that of the geographic analysis unit. While simple, this method does not consider street connectivity, and tends to overestimate the population's access to transit services in areas where man-made or natural barriers such as

valleys, rivers, canals, lakes, fences, and walls prevent pedestrian access to the streets. A more accurate approach is to consider the street segments that are connected to a transit stop and are within walking distances. In (Gomez 1998), dynamic segmentation was used to allocate such street segments to transit stops. Then the number of employment in a TAZ is determined based on if the TAZ is connected to (i.e., intersected by) one of these street segments. Because in this study a detailed land use map was used, portions of the population living within walking distance of a transit stop may also be determined based on their proximity to the access streets. The land use categories considered include high density residential (apartment complexes and row houses), medium density residential, and low density residential.

(3) Modeling of proposed routes or extensions to existing routes to determine projected ridership, study of origins and destinations of home-based work trips, and generation of listings of street names and address ranges that fall within the study area (Dallas Area Rapid Transit, DART).

(4) Identification of areas of high population density and low income and the proportion of families that do not own a car. By comparing these parameters with the transit service provided, areas that need improved transit services may be determined.

Although the data needed may vary depending on applications, treatment of parameters, precision afforded, etc., the core data in transit ridership forecasts are demographic and transit data including:

(1) Household. A household includes all the persons who occupy a housing unit. This is a basic element in demographic analysis. Household data may be aggregated at various levels such as TAZ, census tract, census block groups, and census blocks, which are successively smaller in size. The number and type of dwelling units (low-density single family, medium-density single family, multi-family units, apartments, etc.) may be used to further determine distribution of population within a geographic analysis unit. Table 4.1 shows a sample of housing unit data organized by TAZ. The table gives the number of single-family and multi-family housing units as well as the single-family and multi-family population.

Table 4.1 Sample Data of Types and Number of Housing Units by TAZ

TAZ90	HSFHU93	SFVTU93	SFVHU93	SF-POP93	MFHHU93	MFHVT93	MFVHU93	MF-POP93
1	23	7	5	27	348	109	79	333
2	0	0	0	0	0	0	0	0
3	0	0	0	0	202	63	46	226
4	0	0	0	0	834	261	190	796
5	19	6	4	52	450	141	102	749
6	0	0	0	0	939	293	214	888
7	0	0	0	0	1761	256	180	1780
8	20	3	2	49	1250	181	128	1704
9	330	44	14	840	0	0	0	0
10	0	0	0	0	1100	160	113	1723
11	56	8	6	87	4537	658	465	7379

Source: Miami-Dade County Information Technology Division.

(2) Family Income. Total amount of income of persons 15 years old and over in each family. The relationship of income and travel by transit is often reversely proportioned. Table 4.2 illustrates family income distribution between $15,000 and $150,000 and up by census tracts. The second column is the number of households in a census tract, the 3rd to 9th columns give the percent of the households whose income fall in a particular income range.

Table 4.2 Sample Family Income Data by Census Tract

TRACT	INC_BASE	INC_150KUP	INC100_150	INC_75_100	INC_50_75K	INC_35_50K	INC_25_35K	INC_15_25K
090200	1629	0.680000	2.760000	7.490000	22.219999	23.760000	14.430000	12.830000
090100	1751	1.260000	1.940000	5.540000	16.110001	20.670000	14.960000	16.730000
130200	2212	2.710000	3.390000	5.740000	21.110001	24.410000	12.390000	13.520000
130100	2233	1.520000	1.610000	7.570000	19.840000	21.360001	18.360001	13.170000
090500	2848	0.490000	3.340000	10.850000	23.809999	19.030001	11.480000	14.360000
090400	1604	0.000000	1.430000	2.740000	17.080000	15.400000	18.889999	15.520000
130300	2810	1.710000	4.130000	8.470000	23.590000	19.959999	14.310000	13.670000
130400	3639	0.880000	3.850000	6.130000	20.670000	19.590000	18.440001	15.970000
090300	859	0.000000	2.790000	3.960000	20.840000	21.650000	13.620000	10.940000
130500	4580	3.650000	7.070000	11.030000	22.969999	20.830000	13.230000	9.610000
090600	2818	1.030000	2.950000	6.780000	19.480000	20.299999	14.230000	14.830000
090800	4040	1.490000	4.500000	9.310000	28.170000	26.360001	11.460000	9.780000
090700	3226	0.650000	1.460000	5.640000	28.889999	26.879999	14.140000	12.550000
011500	4087	10.470000	16.520000	16.809999	20.230000	11.230000	8.880000	7.140000
091101	1834	0.490000	4.420000	10.520000	38.169998	25.629999	7.910000	5.620000
050101	3197	1.750000	3.380000	5.910000	25.299999	23.049999	15.300000	10.670000
091002	1857	1.290000	4.250000	13.620000	34.250000	20.889999	8.940000	9.800000
130600	2781	3.380000	5.470000	8.770001	17.120001	21.360001	17.040001	13.880000
050102	3662	0.820000	3.280000	7.320000	19.280001	17.389999	13.000000	17.500000

Source: Environmental System Research Institute, Inc.

(3) Car ownership. The number of cars owned by each family is also reversely proportional to transit use. In Table 4.3, automobile ownership (grouped into categories of no cars, one car, and two or more cars) is given for single-family and multi-family housing units, respectively.

Table 4.3 Partial Auto Ownership Data by TAZ

TAZ90	SF-NOAUTO9	SF-1AUTO93	SF-2+AUTO9	MF-NOAUTO9	MF-1AUTO93	MF-2+AUTO9
1	4	8	4	157	69	13
2	0	0	0	0	0	0
3	0	0	0	92	40	7
4	0	0	0	379	162	32
5	5	5	3	205	87	17
6	0	0	0	427	182	37
7	0	0	0	924	477	104
8	5	11	1	537	378	154
9	36	111	139	0	0	0
10	0	0	0	578	298	64
11	23	20	5	2381	1231	267
12	0	0	0	727	377	81
13	0	0	0	724	142	0

Source: Miami-Dade County Information Technology Division.

(4) Population density. Population density represents the number of people residing per unit area. It best describes the daily trip potential in the origin area. One way to determine the population density is simply to divide the total population in a geographic analysis unit by the area of the unit. Again, more detailed land use information allows a better description of population distribution. Table 4.4 shows a partial attribute table of a land use layer that has detailed land use categories including public and community-owned street rights-of-way, water, parks, varying density single- and multi-familty residential areas, mobile homes, etc. However, in the case that such data is lacking, population is usually assumed to be evenly distributed in the geographic analysis units.

Table 4.4 Land Use Classification in a Partial Land Use GIS Layer Attribute Table

AREA	PERIMETER	LU94	LU94_ID	LU	CATEGORY
3546.129000	362.607900	2	13276	30	Multi-family, Low-density (Under 25 DU/Gross Acre).
38938.310000	989.352200	3	13473	530	Golf courses, public and private.
65568.450000	1369.163000	4	13071	918	Inland water bodies (Lakes, Watercourses) associated with residential developments.
208524.800000	3694.747000	5	13485	530	Golf courses, public and private.
283644.200000	4328.310000	6	13071	918	Inland water bodies (Lakes, Watercourses) associated with residential developments.
39073.560000	1162.889000	7	13360	530	Golf courses, public and private.
346845.300000	5988.043000	8	13048	10	Single-Family, Med-density (2-5 DU/Gross Acre).
9269287.000000	255054.500000	9	6	640	Streets and roads (Including Rights-of-way), except expressways and private drives.
40900.640000	971.949800	10	13182	10	Single-Family, Med-density (2-5 DU/Gross Acre).
7254.538000	335.131500	11	13013	10	Single-Family, Med-density (2-5 DU/Gross Acre).
392513.700000	5935.363000	14	13190	10	Single-Family, Med-density (2-5 DU/Gross Acre).
195050.600000	2739.596000	15	13246	918	Inland water bodies (Lakes, Watercourses) associated with residential developments.
55452.130000	967.098300	16	13439	10	Single-Family, Med-density (2-5 DU/Gross Acre).
484203.100000	4542.952000	23	13257	918	Inland water bodies (Lakes, Watercourses) associated with residential developments.
122112.300000	1762.856000	29	12953	631	Major transmission lines.

Source: Miami-Dade County Information Technology Department.

(5) Employment or employment density. Employment density represents the number of jobs per unit area. Data are found in traffic analysis zones (TAZs) of the Census Transportation Planning Package (CTPP). However, because it is reduced to the TAZ level, which is usually not disaggregated enough for detailed service planning, MPO or regional database is often used. More accurate data may be available commercially regarding the actual locations of businesses and number of employees working at the location. Table 4.5 illustrates a database table that contains employment data at the TAZ level.

Table 4.5 Partial Industrial, Commercial, and Service Employment Data by TAZ

IND-JOBS93	COMM-JOBS9	SERV-JOBS9	TOL-JOB93
139	46	548	733
0	437	8684	9121
0	115	68	183
0	46	146	192
29	59	228	316
29	173	541	743
47	115	619	781
47	208	266	521
0	0	176	176
10	241	1325	1576
75	161	874	1110
36	485	758	1279
70	266	793	1129
47	358	1184	1589
47	231	550	828
0	12	181	193
125	531	912	1568
47	1617	4438	6102
47	693	3152	3892
36	508	1307	1851

Source: Miami-Dade County Information Technology Department.

(6) Trip attractors. Trip attractors include types and locations of employment, retail and recreational facilities, hospitals and medical centers, schools, public libraries, etc. Such data are typically represented as point features in GIS, which may be created from their street addresses by geocoding. Geocoding is a process that converts a locational reference such as a street address or mile post into geographic coordinates. Street addresses may be geocoded since most street data are derived from TIGER/Line files, in which each street segment has a low and high address range. A given street address then may be geocoded by interpolating the address between the low and high numbers for that segment, thus finding a point on that street segment.

(7) Transit data. These include street networks and transit routes, transit service frequency

and coverage, travel times, and transit trip costs. These data will be discussed in more detail in the next section in which GIS data needed for service planning are discussed.

Demographic data are descriptive data and may be obtained from census data, local government data, and variously private sources. A key step in data import is to ensure that the demographic data is tied to the corresponding geographic location and that this location is defined consistently with the other geographic data in the GIS. This is achieved by storing the unique identification number (or a *key* in database terms) of a geographic feature (a point, line, or polygon) in its attribute table or other related database tables. The correct relationship between a geographic feature and its associated data may then be established based on the fact that they share the same key.

4.2 Service Planning/Market Analysis

Service planning involves development and improvement of transit services, i.e., to make changes in routes and schedules, which are aimed at improving effectiveness and efficiency and responding to changes in land uses, travel patterns, and resources. Service planning is a short term, operational, and continuing planning process, and should be done on a route-by-route basis.

Potential service changes may include the following:

- addition of new routes
- increase in frequency of service on a particular route
- increase in span of service
- realignment or extension of existing routes
- change in schedules
- discontinuing of routes or route segments with low ridership
- service adjustment on existing routes to reduce duplication and increase productivity

The key steps in service planning include (Levinson 1984):

(1) reviewing characteristics of the service area - including the physical feasibility of the proposed bus routes;
(2) estimating ridership;
(3) estimating revenues;
(4) simulating bus travel times;
(5) estimating service requirements and costs; and

(6) assessing economic performance.

Bus transit service planning should reflect the specific needs and operating requirements of each urban area. Relative planning factors include data of past operating practices and procedures; the current operating authority and system extent, fare box cost-recovery requirement, land use, population density, and employment features; street patterns; and the availability of off-street rail transit. These factors influence the pattern of bus services and the opportunities for change and expansion (Levinson 1990).

There are many ways that GIS may help in the process of service planning. Some major applications may include the following:

(1) Display of the relationship of a new route to existing routes. Data can either be associated with particular stops or full route or route segments. The common terminal or bus stops may be displayed by using the scheduling and running time data, transfer time where the new route intersects an existing one can be identified. GIS may be used to calculate passenger flows along a route and then display these flows in several formats for both internal and external presentations. Attanucci and Halvorsen (1993) point out that it is much easier to discuss service change proposals with Boards of Directors when armed with easy-to-understand presentations of ridership on the routes in question.

(2) New route design. GIS can help control the duplication of bus routes when designing a new route. This may be achieved by identifying close or overlapping service areas of the new and existing route. Duplication of service areas may be shown by overlaying maps of existing routes and service frequency on population and employment patterns to see the need to modify the existing routes. The purpose is to ensure the transit services are adequately distributed geographically within a service area.

(3) Stimulation of stop spacing, stop siting, and travel time for newly designed routes. The purpose of stimulation is to reduce unnecessary delays or detours. Using stop-by-stop ridership data collected on all routes in a bus system, planners can aggregate data across all routes to determine the most heavily used stops to prioritize the placement of new bus stop shelters. By considering population density, household income, car ownership, destination, and street geometric constraints, the location of bus stops may be obtained.

(4) Level-of-service monitoring. GIS may be used to display the level of service characteristics (such as crosstown, feeder, radial, etc.), and to identify the relationship between service level

and ridership. An analysis of transit level related to use can determine where transit service should be increased to attain a high level of utilization.

(5) Examination of sub-area bus circulation and terminal locations. Many commercial GIS packages provide CAD-like design features which allow the user to sketch-in plans on a base map. CAD drawings and imagery of actual sites may also be linked to street center lines or bus stations/stops. In addition, some GIS support the incorporation of one-way street information, the number of traffic lanes, turning restrictions, and traffic or pedestrian barriers on detailed sub-area maps. These features provide transit planners with the tools to plan bus route circulation and layover patterns in congested urban areas (Attanucci and Halvorsen 1993).

Estimating ridership may involve using sophisticated planning models, or applying simple methods such as estimating population and employment with walking distance of transit facilities using GIS. Maps of selected service areas showing information related to ridership demand such as population, employment, population age, family income, etc., may be displayed and produced.

A potential service planning instrument that may be useful is the regional household survey data collected periodically for urban transportation planning and modeling. Such surveys typically involve family members in a household entering records of their individual trips into a journal, which may include origins and destinations of their trips, intermediate stops, time of travel, travel mode, purpose of trips, reasons for using a particular mode (including public transit), etc. Using GIS, the survey data may be analyzed to identify deficiencies in service configuration in terms of disparity between travel pattern and service route configuration. Although no publications were found by the authors on using GIS to analyze travel survey data for transit service improvement, use of GIS to analyze a general travel pattern has been reported in, e.g., (Shaw and Dass 1996, Sun 1997).

The passage of the 1996 welfare reform bill replaces the Aid to Families with Dependent Children (AFDC) with the Temporary Assistance for Needy Families (TANF). As indicated by its name, TANF is temporary since it provides assistance for no more than 60 months, after which the aid will terminate and the welfare recipient should find employment. This presents great challenges to transit properties because to enable welfare recipients to find jobs, most of whom are dependent on public transportation, they must be able to access employment sites, which are often dispersed in suburbs. To be able to provide adequate transit services to welfare recipients to access employment sites becomes a crucial element in the success of the welfare reform.

GIS may be used to assess mobility problems for welfare recipients. By storing addresses of welfare recipients and addresses of potential employers, transit services may be analyzed to evaluate whether these potential employers are accessible to welfare recipients. In addition, accessibility to child care and job training facilities should also be provided.

Key data for service planning include:

- bus stop, route and running time data;
- the density and distribution of income and age of population;
- the available street pattem with particular regard to the suitability of the bus service, including street segments, intersections, etc.;
- locations of employment, retail and recreational areas, hospitals and medical centers, child care facilities, schools, and job training sites;
- demographics and socioeconomic data including welfare recipients' addresses.

A detailed description of these data is presented below.

Transit stations/stops (access point)

Transit station/stop data are used for route design and analysis of level of service and operation schedules (FTA 1996b). Transit station/stop data are typically stored as a point database. Some transit agencies select points directly from a digital base map. Data may also be collected by using the global positioning system (GPS) technology, with which accurate spatial coordinates may be obtained. However, the data collected may cause significant errors when the data is overlaid onto the representation of the road network which possesses different non-linear errors and resolutions. In some instances, bus stops may be located far away from any streets or transit routes. Such errors, however, may be easily identified using GIS tools (Gomez 1998).

<u>Data Composition</u>

The transit station/stop database catalogs all the locations where transit patrons board or alight a transit vehicle. The data set should include at least the following fields:

- Unique bus stop/access point ID: each access point should contain a unique identifier to which every department can access.

- Location reference: each bus stop requires at least one type of location identifier in order to locate it on a digital base map (FTA 1996b), which may include longitude and latitude or (x, y) coordinates in a system such as State Plane Coordinate System or UTM.

- Intersection: intersection information will be useful to verify bus stop locations and for customer service.

- Connections: routes that utilize the stop/stations. The relationship between transit stops and routes is a many-to-many relationship. In other words, each stop may be utilized by multiple routes, and each route consists of many stops. Tables 4.6 and 4.7 illustrate two different approaches to model this relationship. The database table shown in Table 4.6 uses zero and one to indicate whether a stop serves a particular route. In Table 4.7, the relationship between stops and routes are represented by stop-route pairs. The route direction may be inbound (I), outbound (O), or both (B). Other schemes may be used to designate route directions. This approach is more flexible to allow easy addition and deletion of routes or stops, and supports query on either stops or routes.

Table 4.6 Coding Stop-Route Relationship by Stops

Stop#	X	Y	Intersection	Route_1	Route_1A	Route_2	Route_3	Route_3B	....
1	...	...	...	1	0	0	0	1	...
2	...	...	...	0	0	1	0	1	...
...	...	...	...	...	...	...	...	...	...

Table 4.7 Coding Stop-Route Relationship by Stop-Route Pairs

Stop#	X	Y	Intersection
1	...	...	...
2	...	...	...
3	...	...	...
...	...	...	...

Stop	Route	Route_Dir
1	1	B
1	1A	B
...	...	...
2	1	I
2	3B	O
...	...	...
3	1	I
3	56	B
...	...	...

Additional station/stop data may include type of access control, accessibility, shelter, amenities, etc.

Route

Transit routes make up the transit network. Route data may be used to examine if the existing transit network is consistent with current demands, the scheduling and transfer situation. Bus services should be concentrated in heavily traveled corridors, which have the greater service frequency and route coverage usually provided on approaches to the city center. Route structure should be clear and understandable, and service duplication should be avoided (Levinson 1984).

The route data may be created by attaching street segments between bus stops along a route, using a shortest path algorithm to connect bus stops. It will be more efficient to use a linear referencing system to create the route data. Instead of partially replicating street data since routes have the same location as streets, a linear referencing system allows routes to be defined as having a start point on the street network, and a set of segments that are subsequently defined by a length along a certain street. **Figure 4.2** illustrates a bus route being represented using a linear referencing method.

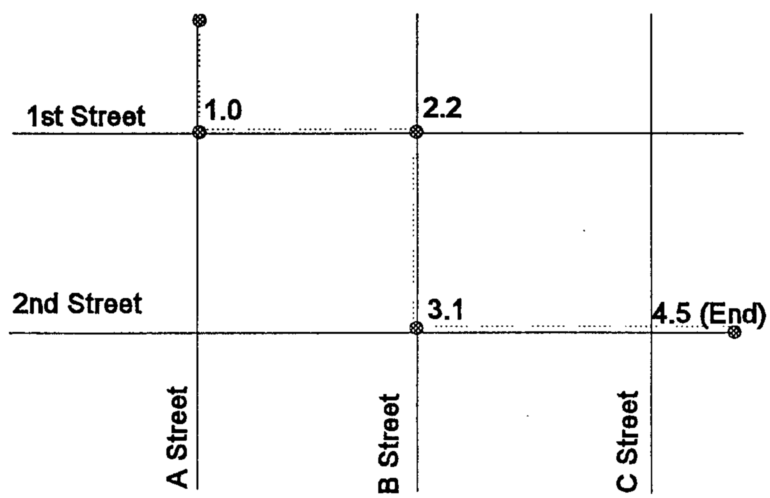

Figure 4.2 A Bus Route Representation in a Linear Referencing System

Data Composition

(1) a point database for each route
(2) a segment table for each route
(3) a network database for the route network

Information should include (FTA 1996b):

- the sequence of bus stops
- bus stop information
- distance/travel time between stops
- time point information

The distance and route direction information may be derived from a digital base map.

Ridership

Ridership demands influence the route structure and service frequency. The ridership is influenced by car ownership, family income, and residential and employment density. In service planning, actual ridership number along each route should be obtained through rider count survey. Ridership may be specified for every route segment and stop of the route network. It may be stored as a database item in the route attribute table or as a separate database table.

<u>Data composition</u>

Ridership data used for service planning should include the number of passengers in each direction of a route during different times of the day. An example of a ridership table is shown in Table 4.8 that contains detailed boarding and alighting data at individual bus stops, along with the route identification, route direction, date, and time.

Table 4.8 Ridership Database

Stop	Route	Route_Dir	Date	Arrv_Time	Dept_Time	AP-Off	AP-On	BP-Off	BP-On
1	1	I	2/21/96	9:30	9:31	0	1	0	0
1	1A	O	3/4/96	17:30	17:32	0	2	1	3
...	...	...	...	...	...	...	...	...	...
2	1	I	2/21/96	9:45	9:47	0	4	0	0
2	3B	O	5/25/96	18:03	18:05	1	6	1	0
...	...	...	...	...	...	...	...	...	...
3	1	I	2/21/96	10:01	10:02	1	4	0	1
...	...	...	...	...	...	...	...	...	...

Running Time

Running time, the amount of time required to traverse a route, is used to check the adherence to the existing schedule. The amount of running time depends upon several factors, such as the type of road, posted speed, the number of passengers boarding and alighting, and the congestion level on the road, etc. Running time may be stored as an impedance of the route. Impedance measures are resistant to movement and are an attribute of arcs and turns.

Data Composition

Running time data include access time, waiting time, travel time, and transfer and departure time.

Street Network

Street network is the transportation network that transit routes operate. It also provides transit users access to transit services. Street network configuration is important when considering transit service accessibility. The traffic condition of the network also affects the transit level of service, therefore the ridership.

Street network can be obtained from several sources. The most common is the TIGER/Line files developed by the U.S. Census Bureau.

Data Composition

Each street segment should consist of street capacity, speed limit, number of lanes and impedances.

Demographic and Socioeconomic Data

Demographic data include population density, number of workers per household, household size, age, and occupation. It also includes household income, auto ownership, floor space, employment types, density, and school enrollment. By locating the young and elder population by tract and comparing the transit service provided (in terms of trips per day per person), areas where the improvement of bus service is needed can be determined. Areas of low income can also be identified. Demographic data have been introduced in the transit ridership forecast in Section 4.1, as well as employment data.

For the purpose of providing mobility to welfare recipients, addresses of welfare recipients may be obtained from state agencies responsible for administering a welfare program. Alternatively, census data may be used to identify residential areas with low income households. Potential employer data may be obtained by collecting information from job advertisements or by studying employer site data to identify those that are likely to hire people with the types of skills of most welfare recipients. Locations of day care facilitie may be obtained from telephone books, and job training sites may be obtained from the state or local agencies that are in charge of job training programs.

4.3 Service/Operations Monitoring

Transit systems should be closely tuned to ridership requirements from a capacity and a cost standpoint. The challenge is to carry passengers comfortably and expeditiously, at a minimum expense to the agency. If service could be performed in equal hourly amounts throughout the day, the task would be simple. Ridership, however, varies throughout a day, a week, and a year due to changes in economy, school operations, weather conditions, and changes in urban development. These traffic variations with space and time call for constant monitoring and adjustment of services.

From the prospective of bus operators, bus trip analysis is a basic component for operation monitoring. Bus trip analysis can provide summary of route-level ridership and running time information, and produce reports on route performance and information of the variation of ridership and environment, which is required to adjust service and provide reliable, punctual and safe service.

Three kinds of tasks should be concerned in bus operations monitoring. The first deals with real-time dispatch and control. Exception reports should be developed for weekly, monthly and quarterly analysis. Finally, route and schedules should be examined periodically to evaluate route usage and performance.

Two important operational measures may be examined by using GIS:

(1) Schedule adherence may be monitored for the same or for different routes for schedule adjustments or evaluating the driver's performance.

(2) Passenger loads as well as the change in the load pattern may be monitored over time.

Using customer boarding and revenue data from the farebox system, Capital District Transportation Authority (CDTA) developed a route performance monitoring system that provides monthly business performance measures on each of the routes (Reily and Guggisberg 1997). The system produces reports each month on average weekday passengers, passengers per hour, revenue to cost ratio, and margin per passenger for each route. These performance measures are disaggregated by service - weekdays, Saturdays and Sundays. Further, time series data on average weekday customer levels are maintained to ascertain trends over time.

With the introduction of the new technology such as automatic vehicle location (AVL) and automatic passenger count (APC), running time and ridership data may be collected at any time and buses will be continuously monitored for schedule adherence and emergency situations. If an agency has implemented APC, it may utilize data such as the number of seats on a bus and the maximum passenger capacity of transit vehicles to determine the suitable fleet size for each route.

Data used in operations monitoring include (Elbadrawi 1996):

- Driver performance
- Vehicle schedule adherence
- Passenger loads
- In-vehicle and off-vehicle security/safety
- Local traffic conditions

Two widely used methods for collecting ridership and schedule adherence data are ride check and point check surveys. In recent years, many transit properties have begun to install and operate AVL on their transit fleet. AVL is also effective for collecting schedule adherence data and ridership data if automatic passenger counters are installed on transit vehicles. In case of ride check surveys, the checker rides selected bus trips and counts the number and times that passengers arrive, alight, board, and leave. Point check surveys are often taken in a few bus stops along the route, in which the checker records the bus number along with arrival and departure times and number of passengers arriving, alighting and getting off.

Continuous ride check data collection protocols based on specified sampling methods are required by the Federal Transit Administration to arrive at estimates of annual systemwide unlinked passenger trips and passenger miles that must be reported under the Section 15 requirements of the Urban Mass Transportation of 1964, as amended (Papacosts 1995). Ride and point checks are also conducted for the purposes of comprehensive operations analysis (COAs), individual route studies, subsystem studies, and in response to localized problems.

In a research conducted by Papacostas (1995), titled "GIS application to the Monitoring of bus operation", the data collecting and reporting requirements related to the compliance of transit properties with Title VI of the Civil Rights Act are presented. They are summarized as follows:

(1) A scaled base map of the transit service area is constructed identifying each census tract or traffic analysis zone (TAZ), major streets and highways, and major activity centers or transit trip generators.

(2) Minority population for each census tract or TAZ in raw numbers and as a percentage of the subarea's total population and all transit routes in the service area is displayed.

(3) A population/racial distribution chart tabulating the actual numbers and percentages of each minority group within each zone or tract may be created.

(4) Five indicators considered to be significant to monitor public transit systems in compliance with Title VI (FTA) include vehicle load, vehicle assignment, vehicle headway, distribution of transit amenities, and transit access.

Locational information provided by an AVL system is useful for operations in several ways besides collecting ridership and schedule adherence data. These include vehicle component monitoring (e.g. engine conditions out of tolerance are flagged and dispatch notified), detailed fare information collection by stops when used in conjunction with an automated fare payment, and improved fleet management with better ability to track transit vehicle movement, early discovery of problems, and quick response to problems that may cause serious service delays or disruptions. For more information about AVL, refer to (FTA 1996c, CUTR 1996, Goeddel 1996, Zhao *et al.* 1997).

Besides the data described in the service planning section (bus stop, route, ridership, running time, street network, and demographic data), the following data are also important in the operation of monitoring applications.

Bus trip

Bus trip is the basic analysis unit for operations monitoring. A bus trip is a productive journey which can be carried out by a vehicle. The information can be obtained through the ride check along with the point check. A bus trip is characterized by a route, a place of origin, a destination, a duration, a cost and a time interval.

Scheduling Data

Scheduling data are overall descriptions of expected and actual travel times on particular routes or portions of a route as represented by transit schedules or timetables. This may also include a description of the fare charged for travel on particular routes. Schedule adherence is one of the service criteria that measures reliability (TRB 1995).

Farebox Data

Description of revenues collected by route, by stop, or by particular types of fare, e.g. revenue from annual passes, monthly regular passes, student passes, senior discounts, transfers, cash, etc.

4.4 Customer Information Services

GIS may be used as a powerful tool to provide the transit users with information about transit services on a digital map that may be used for pre-trip, in-terminal, or in-vehicle planning. Travel information may be delivered interactively via telephone, kiosks at public places, and the Internet. Functions of a customer information system include:

- Monitoring and responding to passenger comments and complaints;
- Disseminating routes and schedule information;
- Disseminating real-time vehicle and route status;
- Disseminating fare information;
- Receiving incoming service requests (for demand responsive service); and
- Disseminating ADA-required mobility information.

GIS applications to customer information service may include:

(1) Query. In-terminal geographic information systems provide schedule updates and transfer information for passengers. This information includes arrival and departure times, information on transfers and connections, information on other regional transportation services, and park and ride facilities. On board of a vehicle, information on routes, schedules, and connecting services are provided via visual displays and annunciations. By specifying the trip origin and destination, information such as route to be used, arrival time at each stop, transfer points, fares, nearest parking lots, ADA accessibility may be obtained.

(2) Automatic Trip Planning (ATP). ATP can provide the customer with the optimal trip plans. When a customer calls, the nearest bus stop is identified and an optimal trip based on shortest travel time or customer's own preference may be obtained. As discussed in (Peng 1995), an ATP system should have an automatic geocoding function to identify the relationship of the origin and destination and the bus routes and stop. If there is more than one route serving a bus stop, some priority measures can be made for the planning process. For more detail refer to (Peng 1995).

For more accurate information, an automatic vehicle location (AVL) system may be used with GIS to provide real-time information. Scheduled and route information may be adjusted according to traffic delays, accidents or emergency situations.

The data needed would be general data of the transit agencies, such as schedule and fare information, transit routes, real-time vehicle data, etc. These data have been described in detail in the previous sections, and will not be repeated here.

4.5 Facility Management

There has been much research on infrastructure management. While for different applications infrastructure may mean different types of physical facilities, they all share the same fundamental goals, objectives, and principles. Facility management involves continuous monitoring, maintenance, and repair of facilities to ensure the efficient use of existing and future resources, safe traffic operations. To make effective decisions regarding facility management with the goal of minimizing costs and maximizing utility of the infrastructure, a transit property has to have accurate information about the condition of the facilities, their past, present, and future usage, and their projected conditions, which are often determined utilizing predictive models. Decisions as to what, where, when, and how maintenance and rehabilitation should be performed need to be made. According to McNeil *et al.* (1992), the five components of the decision process are: (1) data collection and monitoring; (2) impact modeling and application of impact models; (3) strategy selection; (4) strategy implementation; and (5) objective specification and re-evaluation.

Transit facilities include fixed and mobile facilities. Fixed facilities include rail bridges and guideways, underground tunnels, rail storage yards, transit stations, park and ride lots, bus stops, bus shelters, etc., Mobile facilities include transit stop signs, billboards, and maps. Facility management refers to the ability to manage transit facilities and real estates based on several characteristics including location, inventory and conditions.

Inventory data are fundamental in facility management, and GIS may be employed to provide better management of spatially distributed infrastructure data. The main contribution of GIS is its ability to geographically reference the facilities to routes, street segments, and to actual geographic coordinate locations.

The three kinds of facilities in a transit organization that involve geographical referencing of inventories of facilities are as follows (CUTR 1996):

(1) Major vehicle maintenance facilities and dispatch center facilities that have street addresses. They may be represented as points in a GIS.

(2) Bus shelters and bus signs that exist in street rights-of-way that do not have street addresses. They may also be represented as points in a GIS and referenced by either their actual coordinates or a linear referencing system.

(3) Separate rights-of-way for rapid transit facilities. Depending on the type of rights-of-way, they may be represented differently. Bridges used exclusively by transit vehicles may be represented as points while elevated guideways, underground tunnels, and rapid transit alignments may be represented as lines.

The attribute data needed for each inventory of a facility are listed as follows (Laver 1997):

(1) Age. Asset age measured in years. Older ages imply declining physical condition.

(2) Condition data. Conditions reflect the facility usage. A higher rate of usage will increase the decay rate. Actual measurements of the rate of the use vary by facility type:
   - Train miles and/or vehicles per period (for track, guideway, etc.)

- Ridership (for stations)
- Vehicle miles (for vehicles)

(3) Maintenance rate. An increasing rate of maintenance will decrease the deterioration rate. Measurement of the rate of maintenance varied with facility type. Typical measures include the following:

- Annual labor hours (i.e. for direct maintenance staff)
- Annual maintenance expenditures

(4) Rehabilitation history. In general, a rehabilitated facility will have a higher condition rating than an asset of similar type and age but which has not undergone rehabilitation. This may be represented by the date and expenditure of each rehabilitation.

The above data should be used in combination with developing forecasting models, otherwise it will be of little use.

4.6 Transit Safety and Security

Transit safety involves the prevention of accidents. The elements of a transit safety program include employee training, equipment and facility maintenance, vehicle operations, public information, accident record keeping, analysis, and reporting, and establishment of safety goals (Lester 1984). Safety for bus transit also involves vehicle designs that improve the likelihood that an accident will not occur. Features added to improve safety are turn signals on the sides of buses, energy-absorbing (water) bumpers, nonskid floors, adequate lighting, adequate clearance for doors, and visibility of doors and windshields. Safety for rail transit involves considerations of both operations and equipment and facilities (Lester 1984).

Transit safety problems may result from many factors. Some of these factors are a lack of training of transit operators resulting in a poor understanding of safety procedures, equipment defects, poor operating environment, inadequate maintenance, poor accident record keeping and reporting, lack of employee awareness and active participation in safety programs.

GIS may be used to perform analysis of accident data in response to complaints related to individual sites (Mark *et al*, 1990). Using simultaneously the accident data and roadway/guideway design features and operation characteristics, GIS may assist in identifying unsafe factors. For instance, by combining the location and operation data of a bus route, operation procedures may be checked to avoid dangerous corners or intersections, identify inadequate bus-stop distances for boarding, alighting, and merging with traffic, avoid railroad crossings, and adhere to traffic controls.

The data needed for transit safety analysis using GIS may include vehicle operating data such as headway, schedule, bus trip, etc.; facility maintenance data; transit data such as bus stop and route location; street network data; and accident records.

First four kinds of data have been described previously. Therefore, the discussion here focuses on accident data.

In fact, the accident data used for transit safety is of no difference from that for highway safety. Usually the data of the greatest interest include the following items:

(1) Location and direction of travel prior to the accident, of all vehicles contributing to the accident, of all vehicles contributing to the accident, including those stopped or parked.

(2) Time, day of week, and date of accidents.

(3) The general type of accident and manner of accidents.

(4) The movements intended by drivers or pedestrians immediately before the accident (stop to park, to turn left at a specific locations, etc.).

(5) The light condition, weather, and road condition at time of accident.

(6) The type of traffic control affection one or all of the involved traffic units.

(7) The severity of the accident (fatal, injury, or property-damage only).

GIS applications may be developed to improve transit security. GIS can perform incident analysis and reporting. Through incident analysis, crime patterns and security loop holes may be identified. These security problems may then specifically targeted by either improve facility security equipment or adding more police patrol. In addition, security incident information may be used to help employees and customers to avoid places where there is a higher rate of crimes.

The data needed for developing GIS security applications include:

- incident reports. Incident reports should include the date, time, location, nature of the incident, and detailed description of the incident. Incidents that occur on transit properties or occur in areas

- crime statistics in the region and in the immediate areas that have transit properties. Such statistics may be used to determine the security level in an area, and help identify whether or not a particular crime occurred at a transit stop is a reflection of the general crime pattern in the area.

- transit routes and facilities such as stops and stations. The security features including lighting condition, visibility, layout, presence of security personnel, emergency phones, and public address system, etc., should be included.

- police jurisdiction of the area where an incident occurred.

4.7 ADA and Paratransit

GIS may be used to maintain a database of paratransit users, study demand for paratransit services, analyze paratransit users' needs for accessibility to employment, medical, and social services, and assist in scheduling and dispatching of paratransit vehicles. Paratransit users may be address-geocoded in order to maintain a database of customers. This database is useful to locate a paratransit user by retrieving all the relevant information of address, location, and special transportation needs of the passenger. For individual trips, GIS may be used for vehicle routing to determine trip assignments for each vehicle. Records of services including origin, destination, travel time, travel frequencies, etc., may be useful in analyzing travel patterns of paratransit users to determine the best strategies to meet the demand.

Data needed include customer addresses, trip attraction locations, street network, network impendence, transit network, and related service information.

5. TRANSIT GIS DATA SOURCES

Data acquisition and conversion are a significant part of implementing a transit GIS. The effort necessary for data conversion of a GIS can be determined by comparing data available from existing source materials with the data requirements of the GIS application. Data acquisition and conversion are usually both lengthy and expensive, and has traditionally been the most expensive part of implementing a GIS (Glenn et al, 1993). To reduce costs of GIS application development, effort should be made to utilize existing data sources, especially the many free government data. In this section, the data analysis units, which are essential in GIS, and data quality are first discussed. A variety of data sources including public and private sources are reviewed, followed by a review of data collection methods.

5.1 Data Analysis Unit

For the purpose of aggregating data, transportation professionals need to process data by geographic location. Two kinds of data units are commonly used. One is census blocks or tracks, and the other is traffic analysis zones (TAZs).

Census blocks are a data unit used by U.S. Bureau of Census for collecting census data and compiling census statistics. Census blocks are small areas bounded on all sides by visible features such as streets, roads, streams, and railroad tracks, and by invisible boundaries such as city, town, township, and county limits, property lines, and short, imaginary extensions of streets and roads. Census tracts are geographic entities within a county (or statistical equivalent of a county) defined by a committee of local users. Generally, census tracts have a population size between 2500 and 8000 people, with an average of about 4000 people.

TAZ is often used by transportation professionals, and the selection of TAZ zones are often based on several criteria (Meyer et al, 1984):

a. Achieving homogeneous socioeconomic characteristics for each zone's population;
b. Minimizing the number of intrazonal trips;
c. Recognizing physical, political, and historical boundaries;
d. Generating only connected zones and avoiding zones that are completely contained within another zone;
e. Devising a zonal system in which the number of households, population, area, or trips generated and attracted are nearly equal in each zone; and
f. Basing zonal boundaries on census zones.

TAZ data are often based on census data although local government agencies may update the population, school enrollment, and employment data periodically.

5.2 Criteria for Evaluating the Quality of a Data Source

Before introducing data sources, criteria should be set up to evaluate a data source. Montgomery (1993) recommends to evaluate data sources by considering the following factors:

- Accuracy: Degree of conformity with standard or accepted values. Accuracy should be stated and conform to the national mapping standards. A data source with unknown accuracy is not reliable in its quality unless proven otherwise through extensive analysis and validation.

- Completeness: Sources should show most, if not all, specific features. For example, roads should not be missing, and all roads should have a name attached.

- Timeliness: the source map or document should contain the most up-to-date information as possible. It needs to be pointed out that GIS data are often collected and compiled at different times. Therefore, different GIS layers may represent information from different time periods, which may be a potential source of errors in analysis.

- Correctness: The correctness of a data source measures whether it shows information that matched the real world. For instance, a four-lane road should be described as such instead of described as a two-lane road.

- Credibility: Source type credibility is also considered as a data quality factor. Often, credibility is established through synchronization among multiple source types.

- Reliability: If field crews and engineers continually encounter field conditions that are dramatically different from what is shown on their operating maps, those operating maps will be considered an unreliable source of information.

- Convenience: Convenience is often an important factor in assessing data source suitability for conversion purposes. This factor merely deals with how easy it is to locate and use the source for data conversion.

- Precedence: As a rule, source material will show duplicated detail. It is usually preferable to utilize the preceding source.

- Maintainability: The ease of maintenance is also an important factor in selecting a data source.

5.3 Public Data Sources

Public data sources refer to government agencies that produce GIS products. Data from public sources are most commonly used by transit agencies. It also forms the basis of various data products provided by private firms. The cost for public sources is mostly only to recover staff cost and materials, therefore public data sources represent a great resource for low-cost data acquisition.

5.3.1 U.S. Census Bureau Products

5.3.1.1 TIGER/Line File

The Census Bureau's Census TIGER (Topological Integrated Geographic Encoding Reference) System is designed to automate the mapping and related geographic activities required to support census and sample survey programs. The Census Bureau uses the TIGER system to create and maintain a digital cartographic database that covers the entire United States. TIGER/Line is the line network product of the TIGER system. The cartographic base for these line networks is taken from GBF/DIME, and from the USGS 1:100,000-scale national map series. DIME is the Dual Independent Map Encoding file containing geographic coordinates of line segments and census area codes. It was used by the U.S. Bureau of Census for the 1980 census and is now being replaced by TIGER format files. In addition to line segments, TIGER/Line files contain census geographic codes and address ranges for the left and right sides of each street segment in metropolitan areas. Three types of data are contained in TIGER/Line files.

- Line features: roads, railroads, hydrography, miscellaneous transportation features and selected power lines and pipe lines, and political and administrative boundaries;

- Landmark: point landmarks such as schools and churches, area landmarks such as parks and cemeteries, Key Geographic Location (KGL) such as a building;

- Polygon: geographic entity codes for areas used to tabulate the 1990 census statistical data and current geographic areas, locations of area landmarks.

TIGER provides, for the first time, an affordable, digitized map base with which census and other data can be associated (Simkowitz, 1990). It is well suited to be used as a base map for transportation organizations of all types. While it does not possess the absolute accuracy required for engineering purposes such as establishing rights-of-way, the relative accuracy is sufficient for most transportation purposes (Simkowitz, 1990). TIGER/Line file can provide street centerline data for a transit GIS database. The released data files contain separate records for each street segment with the following information:

- Longitude and latitude of starting and ending points
- Street name
- Address ranges on each side of the street (for urban area only)
- Statistic or administrative areas on each side of the street

This data structure does not explicitly provide data on nodes, however, it provide pointers to easily establish street intersections and identify codes.

Some improvements of the TIGER/Line files need to be made to support transit applications. The problems in TIGER/Line files include misspelled and incomplete street names and address ranges, and missing new addresses for roads constructed since the 1990 census. Although the TIGER/LINE files of newer versions have been published after 1990, the changes related to updating the actual street network, street locations or attributes are limited.

5.3.1.2 U.S. Census Statistics

The 1990 census data products, released during 1991-93, are available in a variety of new and traditional media. The Census Bureau released a number of useful summary files from the 1990 census that are keyed to the FIPS code of the census areas.

Summary Tape Files (STF.'s) are database files providing statistics with greater subject-matter detail than printed reports. They also present statistics for some types of areas, such as block groups and blocks, that are not included in the reports. STF1 includes 100 percent population and housing counts and characteristics down to the block level. STF 2 contains 100 percent population and housing characteristic, showing more subject detain than STF 1. STF 3 includes sample population and housing characteristics down to the block level.

5.3.2 USGS Products

5.3.2.1 Digial Line Graph File

Data from the United States Geology Survey (USGS) may also be used to define street networks. USGS offers Digital Line Graphics (DLG) files through the National Digital Cartographic Data Base (NDCDB). The DLG reflects the cartographic information extracted from 1:100,000 and 1:24,000 map sheets for the United States, which are developed through orthophotos, which are aerial photos that have been rectified to eliminate distortions caused by camera imprecision, instability of flying parameters, topography, etc. The most recent set of photos are for the year of 1994.

These maps are extremely good in accuracy and quality, particularly the 1:100,000 sheets, which now cover the entire United States. When the DLG are overlaid against digitized 1:24,000 maps, the majority of data are either directly on or within a pixel of the 1:24,000's information. DLG information can be exchanged and directly pipelined into a GIS using DLG as the exchange format. DLG has the advantage of being topologically consistent and is capable of carrying considerable attribute information as well. DLG data files include information from the USGS planimetric map base categories such as transportation, hydrography, contours, and public land survey boundaries.

DLG files provides the following data:

- Boundaries -- political and administrative boundaries.

- Hydrography -- flowing water, water bodies, wetlands, and related features.

- Manmade Features -- built-up areas and populated places with information on population. State capitals and county seats are also identified.

- Transportation -- major transportation systems collected in three separate subcategories: (1) roads and trails, which include major highways; (2) railroads; and (3) pipelines, transmission lines, and miscellaneous transportation features, which include only airports and the Alaska pipelines.

- U.S. Public Land Survey System -- land grants and subdivisions of the public lands to the township and range level.

The deficiency of USGS DLG files lies in that it doesn't include address information or land use type. The best way to assign addresses to these data would be to use TIGER/Line files and match them with streets identified by DLG files.

5.3.2.2 USGS Land Use/Land Cover Maps

One of the USGS thematic map products is the Land Use and Land Cover (LULC) digital files. The maps are designed so that standard topographic maps at a scale of 1:250,000 can be used as a base for compilation and reproduction. In a few cases, LULC files are designed at a scale of 1:100,000. The l:250,000-scale mapping format is usually a quadrangle unit of l° of latitude by 2° of longitude. The map projection used for LULC covers is the Universal Transverse Mercator projection. The LULC files are vector files; but the raster file format is also avilable (USGS 1990).

The LULC files classify land uses using the Level II categories of the Land Use and Land Cover classification system documented by Anderson *et al.* (l976). The classification system is shown in Table 5.1.

Table 5.1 Land Use and Land Cover Classification System Used by USGS

Level I	Level II
1 Urban or Built-up Land	11 Residential
	12 Commercial and Services
	13 Industrial
	14 Transportation, Communications and Utilities
	15 Industrial and Commercial Complexes
	16 Mixed Urban or Built-up Land
	17 Other Urban or Built-up Land
2 Agricultural Land	21 Cropland and Pasture
	22 Orchards, Groves, Vineyards, Nurseries, and Ornamental Horticultural Areas
	23 Confined Feeding Operations
	24 Other Agricultural Land
3 Rangeland	31 Herbaceous Rangeland
	32 Shrub and Brush Rangeland
	33 Mixed Rangeland
4 Forest Land	41 Deciduous Forest Land
	42 Evergreen Forest Land
	43 Mixed Forest Land
5 Water	51 Streams and Canals
	52 Lakes
	53 Reservoirs
	54 Bays and Estuaries
6 Wetland	61 Forested Wetland
	62 Nonforested Wetland
7 Barren Land	71 Dry Salt Flats
	71 Dry Salt Flats
	72 Beaches

Table 5.1 Land use and Land Cover Classification System Used by USGS (cont.)

Level I	Level II
7 Barren Land	73 Sandy Areas Other than Beaches
	74 Bare Exposed Rock
	75 Strip Mines, Quarries, and Gravel Pits
	76 Transitional Areas
	77 Mixed Barren Land
8 Tundra	81 Shrub and Brush Tundra
	82 Herbaceous Tundra
	83 Bare Ground
	84 Wet Tundra
	85 Mixed Tundra
9 Perennial Snow or Ice	91 Perennial Snowfields
	92 Glaciers

The resolution of LULC covers (or the minimum size of polygons) is four hectares for all Urban or Built-up Land (categories 11-17), Water (51-54), Confined Feeding Operations (23), Other Agricultural Land (24), Strip Mines, Quarries, and Gravel Pits (75) and urban Transitional areas (76). The minimum width of a polygon feature is 200 meters. All other categories of Land Use and Land Cover have a minimum polygon size of 16 hectares and the minimum width of polygon features is 400 meters.

Although coarse data, USGS LULC covers provide low cost, readily available land use data, based on which more detailed land use information may be developed.

5.3.3 Census Transportation Planning Package (CTPP)

The Census Transportation Planning Package (CTPP) is a set of special tabulations of 1990 census data tailored to small geographic units -- traffic analysis zones (TAZs) -- to meet the data needs of transportation professionals. CTPP is developed from the long form of the Decennial Census. The bureau produces a standard set of tabulations based on the specification on the required geographic detail submitted to the bureau by local transportation planning organizations. The data is derived from a 17 percent sample and updated every 10 years.

The CTPP tables contain information of population and household characteristics, worker characteristics and characteristics of Journey-to-Work (JTW). Similarities exist between the CTPP and other Census products, such as STF (Summary Tape File) 3A (JHK & Associate, 1995). The notable difference between CCTP and the census products is that the CTPP is the only Census product that summarizes tables by place of work and by place of residence while the Census Bureau products provide only information by place of residence.

The CTPP can provide demographic and trip characteristics data for transit long-term planning, which include data of household, school enrollment, median household income, number of workers in household, vehicles available by household income, means of transportation and travel time to work. These data are important for studying travel behavior, forecasting transit ridership, and identifying areas that need to improve transit services according to the population density, income data, and travel patterns. However, they do not provide information about trip attractors, and thus cannot be used to determine the characteristics of trip attractions.

5.3.4 National Transit GIS Database

The Federal Transit Administration National Transit Geographic Information System (NTG) is a representative inventory of the public transit assets of the country. The purpose of establishing such a database is to encourage transit agencies to use GIS in their planning, operation, facility management, etc. Bridgewater State College (BSC) began the project in cooperation with the FTA during the summer of 1994. By the Fall of 1996, five hundred and fifty of the nation's fixed route bus services had been built into the GIS route systems.

The minimum data products for each transit agency included in the database are as follows:

- A GIS line database of the streets of the county or counties where the service is located. The street data are extracted from the Bureau of Census 1992 Enhanced TIGER files.

- A condensed file (.cdf) of the TIGER-based street network that takes up less data storage and provides quicker access and display.

- A GIS network (.net) built from the county street database.

- A GIS route system (.rts)of the bus routes, built from selected TIGER street segments as indicated on the transit agency's route maps and schedules.

The databases are likely to be outdated. However, they allow a transit property to jump start its GIS applications with some modifications of the databases instead of having to create the same data from scratch.

5.3.5 Local Government Data Sources

There are a variety of data sources in local government agencies. These data are developed for a specific purpose and can be used as a land base or trip generation purpose in transit GIS application. Typical data sources include the following (Montgomery et al, 1993):

- Community oriented data. These data are available in municipalities and local government agencies, and include property address files, building permit files, demographic information, zoning files, accident and crime statistics, environmental conditions, and information on hazardous materials (Glenn, 1993).

- Assessment oriented data. All counties and incorporated municipalities have property taxes assessed to dwelling units and business. Information may include business, square footage of floor areas, and number of employees, or owned assets.

- Land use data. Some local governments, such as state water management districts, often compiled land use data. These data may be compiled from aerial photos, USGS maps, satellite imagery and local data. The land use data provide land use classification information, such as urban and built-up are, agricultural range, water, transportation, communications, and utilities data. Detailed land use data are also becoming available from county or city governments as they begin to build parcel maps, from which land use maps may be created by aggregating parcels with same zoning code.

- Accounting oriented data.

5.4 Private Sources

There are numerous companies that provide commercial geographic data. Here we briefly introduce some typical data providers that provide demographic, street center line, and boundary file data of the United States. The commercial data sources are more expensive. The costs vary, often determined based on the size of geographic areas or by record (Post *et al.* 1996). However, they are processed and tailored to the needs of the customers. In some cases, the data are also improved in quality, especially attribute data.

5.4.1 ETAK, Inc.

ETAK provides road network coverage used for road navigation for the United States and selected European countries. The information is typically based on 1:24,000 map sheets, is cartographically accurate, topologically consistent and has good quality address geocoding.

Because it has been precleaned with topological algorithms, it can be imported into ARC/INFO quickly. EtakMap Premium maps are maintained on an ongoing basis with updates of roads, road names, addresses, and ZIP Codes.

The information provided by EtakMap Premium includes:

- Population.
- Adress and ZIP Code information.
- Road classification including major interstate highways, limited access and U.S. highways, state highways, arterial, collectors, light duty roads, alleys, ramps, and unpaved or unimproved roads.
- Political and statistical geography.
- Major landmark point features in selected areas, including: city centers, major airports, colleges and universities, regional shopping centers, major sports facilities, tourist attractions, major museums, concert halls and auditoriums, and major zoos.
- Urban one-way street encoding in selected areas.

5.4.2 Navigation Technologies

Navigation Technologies provide street-level databases in the following categories:

- Road network geometry
- Connectivity of roadway
- Roadway names including alias/alternate names
- Block-by-block address ranges (both sides of the street)
- Postal codes and political entities+ names for roadways
- Vanity address handling (such as IBM Plaza, United Nations Plaza)
- Roadway classifications, including toll roads, limited access roads, major arteries, main streets, residential streets, and low access roads
- Roadway characteristics, including one-way information, turn restrictions and dividedness
- Overpasses and underpasses, recorded in a relative elevation level representation

5.4.3 Geographic Data Technology, Inc.

Geographic Data Technology, Inc. (GDT) develops and sells geographic digital street network and boundary databases, geocoding software, and geographic services.

GDT continuously updates a nationwide street database, which forms the foundation for its cartographic database products, including Dynamap/1000, Dynamap/2000r and 5-Digit ZIP Code boundaries. GDT's premier nationwide street centerline database includes over 13.5 million addressed segments, improved street classification codes, logical layering for hydrology, and landmarks. GDT's fully realigned nationwide street database contains more address information, census tract and county boundaries. GDT also produces a comprehensive database for geocoding through the addition of new streets and postal information in the ongoing updating process.

5.4.4 American Digital Cartography, Inc.

American Digital Cartography (ADC) offers a range of customized mapping products in a variety of GIS software vendor formats. ADC products are divided into two major categories: Foundation Series Data and Project Specific Data. These data include 1:100,000-scale DLG data from USGS sources and data from U.S. Census Bureau. For more unique mapping projects, ADC provides data configured specifically for each application.

5.4.5 America.dbf, Inc.

America.dbf provides STFXtract Census Demographics on CD-ROM with Transportation Attributes and National Street and Boundary Digital Map Set on CD-ROM. The source of the data is TIGER/Line.

America.dbf offers the complete 1990 Census Demographics STF 1A and 3A datasets in one package that is easy to use. This package allows data lookups by state, county, city, place, census tract, and census block group. Selection of demographic fields are aided by plain English labels, and all 4,800+ fields of the STF census files are included.

Another feature of STFXtract is that any set of demographic fields or any set of geographic selections used in a data extraction can be saved and re-used. This means that one can rapidly create the same extraction on any number of geographic locations in order to compare and contrast the demographics of these locations.

5.4.6 American Business Information, Inc.

American Business Information, Inc., specializes in providing business information. The information that may be used by transit properties is business name, location, mailing address, type of business, number of employees, professional specialities, and latitude and longitude coordinates. Such information may be used for long term transit planning, service planning, market analysis, and marketing.

5.4.7 Business Location Research

Business Location Research (BLR) provides files that are continually updated. BLR maintains over 30 unique databases, including StreetNetwork 6.0, TrafficVolumes 3.0, MajorRoads, HighwayNetwork, Telco boundaries, and demographics data.

5.4.8 Dun & Bradstreet

Dun & Bradstreet (D & B) is a private data source about businesses and their attributes. This source is used by many businesses to acquire credit information on another firm. The data consists of the name of the firm, address, latitude/ longitude, number of employees by location, by SIC code and a contact person (Post, 1996). This source is developed by phone calls to businesses and checked through a variety of sources including banks, utility records, Secretaries of State, and the courts. Addresses are based on actual employment location, and data are available for employers with as few as one employee.

5.5 Date Conversion Ability in Popular GIS Software

Data conversion involves transforming data from an existing format, such as TIGER or USGS format, to the digital format that is required by a specific GIS software being used. Converting publicly available data files is a quick way to start a transit GIS application because these data files consist of basic information, such as streets and other point, line and polygon features, as well as addresses. Different GIS software packages have different data conversion capabilities. The software packages examined here are those widely used in transportation including Arc/Info (ESRI), MGE (Intergraph), GeoMedia (Intergraph), MicroStation GeoGraphics (Bentley), AtlasGIS (ESRI), TransCAD (Caliper), GISPlus (Caliper), and MapInfo (MapInfo Corp.).

Table 5.2 Summary of Data Conversion Capability in GIS Software

Software	TIGER/Line	DIME	USGS DLG	ETAK	IMAGE	DXF
Arc/Info	yes	yes	yes	yes	yes	yes
MGE	yes	yes	yes	yes	yes	yes
GeoMedia	yes	yes	yes	yes	yes	yes
Atlas GIS	yes	yes	yes	yes	yes	yes
Geographics	yes	yes	yes	yes	yes	yes
TransCAD	yes	no	no	yes	yes	yes
MapInfo	no	no	no	no	no	yes

5.5.1 Arc/Info (ESRI)

Currently, PC Data Conversion, one of the specialized Arc/Info products, provides the ability to exchange data between PC Arc/Info coverages and one of seven vector formats or one of seven grid cell formats. The file formats supported by PC Data Conversion are listed below.

- TIGER
- DIME
- USGS DLG
- DXF: AutoCAD ASCII Drawing Interchange format file created by AutoCAD to transfer AutoCAD drawings.
- ETAK: ETAK MapBase files are similar to DIME format files. Each record represents a single linear feature with addresses and census information.
- ATLAS: ATLAS*GRAPHICS export format file containing data extracted from Strategic Mapping's ATLAS*GRAPHICS desktop mapping package. These data may consist of points, lines and/or closed loop polygons and primary and secondary names.
- IGES: Initial Graphics Exchange Standard published by the US Department of Commerce.
- MOSS: Map Overlay Statistical System format file readable by the US Department of Interior's MOSS public domain GIS.

Arc/Info also supports a variety of grid file formats.

5.5.2 MGE & Geomedia

MGE (Modular Geographic Environment) and Geomedia are the products of Intergraph Inc. MGE

GIS Translators/US (MGT-US) can translate industry-standard, publicly available DLG, TIGER, Public Law data (PL 94-171), ETAK MapBase, and GBF/DIME data files into Intergraph's Modular GIS Environment (MGE).

Image Translators is a product that support SPOT, SPOTView, AVHRR, ERS, EOSAT, and USGS DOQ image file formats. Running on Windows NT or Windows 95, Image Translators can translate image data to TIFF or Intergraph standard raster format, and may also be used to copy images from tape-to-tape or tape-to-hard drive. Additionally, it can register images to a MicroStation design file, given sufficient geographic header information, and create image overviews for display while translating files.

GeoMedia supports MGE, FRAMME, Arc/Info, and MicroStation data in a single environment. This means that within one window, several disparate data types can be viewed or used as input to a spatial query.

5.5.3 Atlas GIS (ESRI)

Atlas GIS is a 16-bit Windows application and desktop mapping program designed for the business geographic market, but has also been used for transportation applications.

Atlas Import/Export (IE) imports geographic and associated attribute information into Atlas GIS. The file formats supported by IE include:

- TIGER: TIGER line files F41 (1990) and F51 (1992) series.
- USGS DLG: DLG optional level 3 format only. It doesn't support DEM data.
- Dime: GBF/DIME file
- Arc/Info: uncompressed, Aro/Info Interchange files
- Etak: Etak mapbase files
- MapInfo: MapInfo for DOS Boundary Interchange file, MapInfo for Windows, Macintosh, and UNIX Interchange files, and MapInfo for DOS Map Interchange files

5.5.4 GeoGraphics (Bentley)

Bentley Systems, Inc., provides engineering software products. MicroStation Geographic is a mapping toolbox that combines MicroStation's data capture and editing tools with a database interface and a spatial analysis engine.

A variety of data in IGES,DXF, DWG, and CGM formats may be imported into MicroStation Geographic using import tools.

5.5.5 TransCAD/GISplus (Caliper)

Caliper is the developer of the MAPTITUDE, GISplus, and TransCAD commercial geographic information and analysis software packages.

GISplus is a MS-Window-based GIS. TransCAD is a GIS for transportation and logistics. TransCAD can directly import AutoCAD DXF, ETAK Streets, TIGER/line files Arc/Info E00 and Ungenerate and ArcView Shape files. The current version does not support USGS, DLG and DIME files directly, which will be supported in the future according to the company. TransCAD also supports other file formats such as AtlasGIS, MapInfo, TIFF, digital orthophotos, EMME/2, TranPLAN and MINUTP.

5.5.6 MapInfo (MapInfo Corp.)

MapInfo Desktop and MapInfo Professional are the products of MapInfo Corporation. MapInfo Desktop does not have the ability for importing external files into the product. MapInfo users can only import DXF files or MapInfo Interchange Format(MIF) files. To import files of other formats into MapInfo, a translation should be conducted first. MapInfo does have translators available that will translate ArcInfo, Atlas and Shapefiles from ArcInfo into MapInfo format. Third part vendors also sell translators available that can convert other formats into MapInfo.

6. CONCLUSIONS

In this report, typical transit GIS applications are reviewed. The data needs for various transit GIS applications are identified. The sources from which GIS data may be obtained are also identified and described. While developing transit GIS applications is not an easy task, and some basic training and understanding of GIS are necessary, it does not necessarily demand a full-fledged GIS department to implement some basic and simple applications. These may include demographic analysis of the service areas, analysis of route performance, maintaining records of paratransit users, providing travel information with the assistance of GIS, etc. To make transit GIS applications low cost, it is important to select a user-friendly GIS program, preferably desktop GIS, for the end users. Even for a large transit property where GIS expertise is available, desktop GIS will allow transit professionals to easily access the GIS data and use the data in their daily work.

Another way to keep the cost down is to select the right data to use. The right data should have a locational accuracy adequate for the intended applications and the needed attribute information with good accuracy (e.g. correct street names, correct travel time). Before any field data are collected, existing data sources should be exhausted, such as the National Transit GIS Database and the TIGER/LINE files. It is likely that little or no field data collection needs to be done for implementing simple applications as mentioned before.

In addition to the quantity, quality, and availability of data, data quality is another important issue. Because the cost associated with data collection is high, data sharing among different local government organizations including transit agencies is highly desirable. The difficulty, however, of sharing data lies in the lack of data collection coordination and, in some cases, different data needs of the departments that share the data. Even if data are collected by the transit agency for transit purposes, data quality standards and collection procedures should be carefully designed to ensure good spatial analysis results and maximum utility of the data for different purposes by different divisions or departments.

An important data issue is the maintenance of the data. First, there is always a lack of resources to constantly update GIS data or correct the data. Data may be rather old and still have to be used for current applications. Sometimes it may happen that one GIS layer is more current than another and the two do not agree. For instance, a bus route layer may be current, but the bus stop layer may be several years old, resulting in some of the bus stops missing or some other not being served by any routes. The users should always be familiar with the data, including the source, data collection method, date of create, accuracy, projection and coordinate system used, update history, etc.

Second, GIS data maintenance practice is not well established. There is often a lack of formal procedures for data maintenance. Data are sometimes maintained by individuals who develop them and/or use them. Due to change in personnel, the data history may be lost. It is therefore important to document all changes made to the GIS data, together with the original information describing the data (often called meta data), including both locational data and attribute data. Such documentation allows past changes to be traced and provides useful information to both database administrators and users.

To give transit properties easy access to information related to transit GIS applications, data sources, and software, the establishment of a central information clearing house may prove effective. This information center can distribute information through the web by providing hot links to various Internet sites that contain relevant information. In addition, this center can provide basic training to transit professionals so that they have a basic understanding of the usefulness of GIS.

REFERENCES

Abkowitz, M., Walsh S., Hauser, E., and Minor, L. (1990). "Adaptation of Geographic Information Systems to Highway Management", *Journal of Transportation Engineering*, Vol. 116, No. 3, pp. 310-327.

Anderson, J.R., Hardy, E.E., Roach, J.T., and Witmer, R.E. (l976). *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*, U.S. Geological Survey Professional Paper 964, 28 p.

Attanucci, J.P. and Halvorsen, R. (1993). "What GIS Can Do for Transit Planning", *Proceedings of the Symposium on Operations and Service Planning*, December 8-10, http://www.fta.dot.gov/library/technology/symops/Attanucc.htm.

Austin K. (1994). "An Advanced Computer System for Road Safety Analysis", *Journal of Highway and Transportation*, Vol. 41, No. 11, Institution of Highways and Transportation & HTTA, London, pp. 16-19.

Basile, R.E., Teneyck, T. and Pietropola, A.J. (1991) "Geographic Information System for the Pennsylvania Department of Transportation", *Transportation Research Record*, No. 1305. Transportation Research Board, National Research Council, Washington, D.C., pp. 118-122.

Black, A. (1995). *Urban mass Transportation Planning*, McGraw-Hill.

Brubaker, M. M. (1997). "Transportation Data Issues in the San Francisco Bay Area with National Applications and Recommendations", presented at the 1997 Transportation Research Board Annual Meeting, Washington, D.C.

Business Location Research (1997). *1997 Data Catalog*, Business Location Research, Tucson, Arizona.

Caliper Corporation, *TransCAD User Guide*, version 3.0 for windows.

CTAA (1998). *Access to Jobs - A Guide to Innovative Practices in Welfare-to-Work Transportation*, Technical Report, Community Transportation Association of America, Washington, D.C.

Culp, L. (1997). "Short Range Transit Planning and marketing Using Desktop Geographic Information Systems", San Diego Association of Governments, http://www.fta.dot.gov/fta/library/

planning/CULP/CULP.html.

CUTR (1995). *Transit Service Performance Analysis and Monitoring Process*, Final Report, Center for Urban Transportation Research, University of South Florida, Tampa, FL.

CUTR (1996). "Case Study," in *Proceedings of a Conference on GIS in transit*, Center for Urban Transportation Research, College of Engineering , University of South Florida,Tampa, FL, August 13 - 15, 1995.

Dueker, K.J. and Vrana, R. (1990). "Geographic Information Systems in Urban Public Transportation", *Proceedings of the 1990 GIS for Transportation Symposium*, AASHTO, Washington D.C. pp. 113-121.

ESRI (1994a). *PC Aro/Info User Guide*, Environmental System Research Institute, Inc., Redland, CA.

ESRI (1994b). *PC Network*, Environmental System Research Institute, Inc., Redland, CA.

FTA (1996). *Manual of FTA National Transit GIS*, http://www.fat.dot.gov/fta/library/technology/GIS/ntgistds/NTGISTDS.htm

FTA (1996a). *Data Standards, Guidelines, and Recommended Practices*, Report DTRS57-95-P-80861, Federal Transit Administration, Washington, D.C.

FTA (1996b). *The Standard of National Transit Geographic Information System*, Federal Transit Administration, U.S. Department of Transportation, Washington, D.C.

FTA (1996c). *Fleet Management*, Federal Transit Administration, http://www.fta.gov/fta/library/technology/APTS/update/CHAP2.HTM.

Goeddel, D. (1996). *Benefits Assessment of Advanced Public Transportation Systems*, Technical Report DOT-VNTSC-FTA-96-7, John A. Volpe Transportation Systems Center, U.S. Department of Transportaiton, Cambridge, MA.

Goh, P. C. (1993). "Traffic Accident Analysis Using Geoprocessing Techniques", *Poad & Transport Research*, Vol. 2, No. 2, pp. 76-85, Australia.

Gomez, A. (1998). *Using GIS to Improve Estimates of Population and Employment to Transit Services*, M.S. Thesis, Department of Civil and Environmental Engineering, Florida International

University, Miami, FL.

Hall, R. W., Chellator, E.L. and Yim, Y. "Evaluation of Transit Telephone Information at SCRTD," *Transportation Research Record,* No. 1451, Transportation Research Board, National Research Council, Washington, D.C.

Herman, L.J. and Thatcher, R.H, (1995). *Using Geographic Information Systems for Integrated Transit Service Delivery: Broward County Florida Transit Options Project.* Report 95-0056, Project Action, http://www.projectaction.org.

Hill, C. (1995). *Information Technologies - State-of-the-Art Applications for Transit Technologies,* TCRP project G-1.

Hunt, D.T., Still, S.E., Carroll, J.D. and Kruse, A.O. (1986) "a Geographic Model for Bus Service Planning and Marketing", *Transportation Research Record,* No. 1051, Transportation Research Board, National Research Council, Washington, D.C., pp. 1-12.

JHK & Associates (1995). *CTPP Handbook,* FHWA-PD-95-019, JHK & Associates, Orlando, FL.

Kihl M. (1993). "Advanced Vehicle Location System for Paratransit in Iowa", *Proceedings of the IEEE-IEE Vehicle Navigation and Information Systems Conference,* Ottawa, Ontario, Oct. 12-15, 1993, pp.173-183.

Kikuchi, S., Aneja, S., Chakroborty, P., Hofmann, J., Machida, M., and Perincherry, V. (1994), *Advanced Traveler Aid Systems for Public Transportation: the Intelligent Transit Mobility Systems (ITMS),* Final Report DOT-T-95-07, U.S. Department of Transportation, Washington, D.C.

LaMacchia, R.A. (1990) "The TIGER File and Transportation Issues", *Proceedings of the 1990 Geographic Information Systems (GIS) for Transportation Symposium,* American Association of State Highway and Transportation Officials and US Department of Transportation Federal Highway Administration, San Antonio, Texas, March 14-16, pp77-87.

Laver, R.S. (1997) "The Estimation of Transit Asset Physical Conditions", prepared for Transportation Research Board, TRB Annual Meeting, Washington D.C.

Lester, H. (1984). "Public Transportation Security and Safety", in *Public Transportation,* edited by Gray, G. E. and Hoel, L.A., Prentice Hall, Englewood Cliffs, New Jersey.

Levinson, H.S. (1990). "System and Service Planning," *Public Transportation,* Prentice Hall, edited by Gray, G. E. and Hoel, L.A., Prentice Hall, Englewood Cliffs, New Jersey.

Mayhew, G.L. and Long, D.S. (1993) "Contribution of Vehicle Location Technology to Transit Headway Minimization", *Proceddings of the 43rd Ieee Vehiclear Technology Conference: Personal Communication Freedom Through Wireless Technology,* May 18-20, Meadowlands Hilton, Secaucus, N.J. pp. 356-350.

McNeil, S., Markow, M., Neumann, L., Ordway, J., and Uzarski, D. (1992). " Emerging Issues In Transportation facility Management", *Journal of Transportation Engineering,* Vol 118, No.4, pp. 477-495.

Meyer, M. and Miller, E.J. (1984). *Transportation Planning: a Decision-Oriented Approach,* MaGraw-Hill.

Martin, M.R. (1991). "Geographic Information Systems for Highway/Sign Maintenance", *Microcomputers in Transportation*, Proceedings of the 4[th] *International Conference on Microcomputers in Transportation*, Baltimore, MD, pp. 72-82.

Montgomery, G. E. and Schuch, H.C. (1993). *GIS Data Conversion Handbook*, GIS World, Inc., Fort Collins, Colorado.

Moyer, D.D. (1995). *Proceedings of Geographic Information Systems for Transportation Symposium (GIS-T) 95*, edited by moyer, D.D., Sparks, Nevada, April 3 - 5, 1995.

O'Neill, W.A. (1994). "a Comparison of Accuracy and Cost of Attribute Data in Analysis of Transit Service Areas Using GIS", *Journal of Advanced Transportation*, Vol. 29, No. 3, pp. 299-320.

O'Neil, W.A., Ramsey, R.D. and Chou, J. (1992). Analysis of Transit Service Areas Using Geographic Infomation Systems, *Transportation Research Record* No. 1364, Transportation Research Board, National Research Council, Washington, D.C., pp. 131-138.

Papacostas, C.S. (1995), "GIS Application to the Monitoring of Bus Operations", University of Hawaii, on line paper, http://www.eng.hawaii.edu/~csp/Mygis/busgis.html.

Peled A. and Hakkert A.S. (1993). "PC-oriented GIS Application for Road Safety Analysis and Management", *Traffic Engineering & Control*, Vol. 34, No. 7-8, London, pp. 355-361.

Peng, P., Groff, J.N. and Dueker, K.J. (1995). "An Enterprise GIS Database Design for Transit

Applications", *Proceedings of GIS for Transportation Symposium,* AASHTO, Washington, D.C., pp. 147-164.

Peng, Z., Kenneth, J.D., Strathman, J. and Hopper, A. (1997). "A Simultaneous Route-Level Transit Patronage Model: Demand, Supply, and Inter-Route Relationship," *Transportation* 24: 159-181, Kluwer Academic Publishers.

Post, Buckey, Schuh & Jernigan, Inc.(1996). *Draft Technical Report No.1 Phase I: Data Development and Data Source Inventory,* Fsutms ZDATA2 Development Process Study, State Project No. 99004-1637, WPI No. 4590239, Post, Buckey, Schuh & Jernigan, Inc.

Prastacos, P. (1996). "Integrating GIS Technology in Urban Transportation Planning and Modeling", *Transportation Research Record,* No. 1305, Transportation Research Board, National Research Council, Washington, D.C.

Racca, D. (1995). "Standards for GIS-T", *Proceedings of Geographic Information Systems for Transportation Symposium (GIS-T) '95,* Moyer, D.D. (ed.), Sparks, Nevada, April 3 - 5.

Reilly, J.M. and Guggisberg, T.G. (1997), "Use of Geographic Information Systems for Transit Service Performance Measurement", presented at the 1997 TRB Annual Meeting, Washington, D.C.

Schweiger, C.L. (1991). *Current Use of Geographic Information Systems in Transit Planning,* Final Report, Federal Transit Administration, U.S. Department of Transportation, Washington, D.C.

Schweiger, C.L. (1992). "Current Use of Geographic Information Systems in Transit Planning," *Public Transit: Management, Options, and Planning and Development, Transportation Research Record,* No. 1349, pp. 93-106, Transportation Research Board, National Research Council, Washington, D.C.

Siderelis, C.K. And Dildine, G.S. (1990). "Using GIS for Multi-agency Planning and Review of Transportation Projects", *Proceedings of the 1990 Geographic Information Systems (GIS) for Transportation Symposium,* American Association of State Highway and Transportation Officials and US Department of Transportation Federal Highway Administration, San Antonio, Texas, March 14-16, pp. 103-112.

Simkowitz, H.J. (1991). "TIGER in Action: Making Effective Use of the Census Bureau's TIGER Files", *Proceedings of the 1990 Geographic Information Systems (GIS) for Transportation Symposium,* American Association of State Highway and Transportation Officials and US

Department of Transportation Federal Highway Administration, San Antonio, Texas, March 14-16, 1990, pp. 89-102.

Sriram P and Ardeshir F. (1995). "Knowledge-Based Geographic Information System for Safety Analysis at Rail-Highway Grade Crossings", *Transportation Research Record,* No. 1497, pp. 91-100.

Sun, X., Wilmot, C.G., and Kasturi, T. (1997). "Household Travel, Household Characteristics, and Land Use: An Empirical Study from the 1994 Portland Travel Study," *Proceedings of the 77th TRB Annual Meeting*, Transportation Research Board, January 11-15, Washington, D.C.

TIGER/Line Files, 1994, technical documentation, Bureau of the Census, U.S. Department of Commerce, Washington, D.C.

TRB *(1994). Management Information Systems*, Synthesis of Transit Practice 5, Transportation Research Board, National Research Council, Washington, D.C.

TRB (1995). *Bus Route Evaluation Standards: a Synthesis of Transit Practice*, TCRP Synthesis 10, Transportation Research Board, National Research Council, Washington, D.C.

U.S. DOT (1994), *New Approaches to Travel Forecasting Models: a Synthesis of Four Research Proposals*, Final Report, DOT-T-94-15, U.S. Department of Transportation, Washington, D.C.

U.S. Department of the Interior (1995). *Data User Guide 3: 1:2,000,000 Digital Line Graphics*, (CD ROM), U.S. Geological Survey.

USGS (1990). *Land Use and Land Cover Digital Data From 1:250,000- and 1:100,000-scale Maps - Data Users Guide 4*, U.S. Geological Survey, Reston, Virginia.

Highway Research Board (1970). *Use of Census Data In Urban Transportation Planning*, Report of a Conference, July 9-10.

Zhao, F. and Shen, L.D. (1997), *GIS Applications for Public Transportation Management*, Technical Report, Lehman Center for Transportation Research, Florida International University, Miami, FL.