CHAPTER - 2

DEVELOPMENT OF GIS DATABASE FOR STUDY AREA
CHAPTER-2

DEVELOPMENT OF GIS DATABASE FOR STUDY AREA

2.0 GENERAL

Transportation organizations constantly need to maintain, access, and share information related to transportation system. GIS is increasingly used to capture, assemble and disseminate much of this information (Adam et al. 2001). The data is required for planning, design, usage and management of the transport system effectively. The organizations of transport system stems from movement of peoples i.e. passengers and goods to increase mobility and usability. The physical routes (e.g. road, highway, and rail) are provided by Government organizations, enforcement by another organization (e.g. accidents reporting and speed control by police department), operation and management by the next one (e.g. SRTCs and private operators) and used by multiple users. This generates large volume of data continuously at different locations, time and organizations. GIS-T application covers much of the board scope of transportation. Transportation analysis and decision makers are using GIS tools in infrastructure planning, design, and management, public transit planning operations, traffic analysis and control, transportation safety analysis, environmental impact assessment, hazards mitigation and configuring and managing complex logistics system, just to name a few application domains (Miller and Shaw 2001). Intelligent Transportation System (ITS), including services such as Intelligent Vehicle Highway System (IVHS) and Automatic Vehicle Location Systems (AVLS), are particularly ambitious integration of GIS and communication technologies to a wide variety of transportation services (Souleyrette and Strauss 1999, Waters 1999).

Therefore, the database development and design for transportation system should recognize this and allow efficient access to data from outside sources. The converse to this is also true, since much of the decision making on these issues takes place outside the transport sector. The developed database provides efficient distribution of data to these sources as well. Thus database development should take into account all
these considerations. The database developed will have wide ranging applications within
the organization and across the organization in long run.

The database management system for such a wide ranging application area
has stretched the traditional database technology and incorporated Relational Database
Management System (RDBMS), Geo-relational models, Structured Query
Language(SQL), Object Orientation(OO) and Object Oriented Data Base Management
System (OOODBMS).

Within the organization the transport related data files are totally unrelated
to each other, duplicative and inconsistent. Because, transport related data always have
spatial component, the most natural way to associate elements from different datasets is
through consistent referencing system. Geo-referencing is relating all spatial
measurements to a spatial referencing system. GIS technology provides the core frame
work for an integrated highway information system (Simkowitz, 1989, Singh 1999). The
utilization capability of a GIS database depends upon how it is organized. In database
terminology, the ways of representing data are known as data models. The GIS data
models to present real world data must inter relate the elements of highway transport. In
this chapter, therefore the framework to organize the GIS database models is discussed
and database creation is presented.

2.1 FRAMEWORK FOR ORGANIZING GIS DATABASE
(GIS DATABASE MODELING)

Real world features exist in two basic forms: objects and
phenomena (Lo and Yeung 2006, Burrough and McDonell, 2000). Objects are discrete
and definite, such as building, highway, cities, and national parks. Phenomena are
distributed continuously over a large area, such as terrain, temperature, rainfall, noise
levels and other environmental indices. Geographic data depict the real world in these
two distinct approaches of representing the real world in geographic databases: the object
based model and the field based model (Fig. 2.1) (Goodchild, 1992, Wang and Howarth,
1994, Lo and Yeung, 2006). The final models adopted in majority of GIS softwares are
SAFE AND EFFICIENT REGIONAL ROAD USERS’ MOVEMENT FOR GUJARAT STATE IN GIS ENVIRONMENT

vector model at graphic level representation. GIS are concentrated with one of the broadest information base conceivable (Milne et al. 1993). It needs huge graphic and attributes data in order to function. The GIS database for transportation is no exception to it. Transportation database are wide, voluminous and diverse in nature, hence there is need for a structured approach to database design and development. The ideal database development strategy for GIS-T is characterized as first top down, then bottom up. The top down part has to do with designing and implementing the required databases and bottom up part is to do with the end user of the database to realize particular applications (Singh, 1999).

2.1.1 OBJECTIVE OF DATABASE DEVELOPMENT

The title of the thesis is “Safe and Efficient Regional Road Users’ Movement for Gujarat State in GIS Environment “. The movement of Road Users is on physical network of highways/ bus transit routes with different attributes by using various modes having different origin and destination. The database is for management of resource, facility, area or region. The database development gives base year transportation scenario.

2.1.2 SPATIAL DATABASE

The spatial data are characterized by information about position, connection with other features and details of non-spatial characteristics. (Burroughand Masser, 1998; Department of Environment, 1987). The spatial database is a map. The spatial data attribute can be metrical or topological. Metrical attributes include position, shape and size that can be expressed by spatial coordinates. Topological attributes are those describing characteristic such as connectivity and adjacency that are invariant under distortions and change of scale. Spatial entities are conventionally divided into points, lines, polygons and surfaces. Polygons are the most frequently encoded features in GIS (Burrough and McDonell 2000).
FIG. 2.1 APPROACHES TO MODELING OF REAL WORLD
(SOURCE: LO AND YEUNG, 2006)
The computer representation of the point, lines, polygons and surface entities are in the form of grid (raster) or in the form of vector data (Peuquet and Bascatow, 1984, Aronoff, 1991). Spatial data are two or three dimensional coordinate of points (nodes), lines (arcs), or areas (polygons) representing one aspect of geographic reality (coverage). Descriptive data, on the other hand, refers to the features or attributes of these points, lines or areas (Huxhold, 1991). The geocoded spatial data defines objects that have an orientation and relationship in two or three dimensional space. Each object classified and georeferenced are related according to rules of mathematical topology. Creating topological relationship allows faster data processing, and also allows for performing analysis functions such as route finding, area aggregation and overlaying geographic features (Huxhold, 1991).

The complete GIS have the following major elements (Marble and Amundson 1988; Linden, 1990; El.-Gafy, 2005). The GIS and its concepts are presented in Fig. 2.2. Some of the important aspects of spatial database development are described in subsequent paragraphs.

FIG. 2.2 MAIN CONCEPTS OF A GIS [SOURCE: LINDEN (1990)]

25
2.1.2.1 GEOGRAPHIC SCALE

The representation and storage of information in GIS are separate things. The spatial data stored in GIS can be viewed at any scale. The *zoom in* and *zoom out* functions available in GIS software, often lead novice users to think that map scale is an unimportant subject in GIS.

Transportation application can cover a wide range of geographic scales. These vary from a global scale, from global transportation logistics analysis to an engineering study of the highway interchange. When we design and populate GIS database to meet the need of various transportation applications, consideration must be given to the positional accuracy level appropriate to the application at hand. Hence when discussing map scale and travel surveys, or almost any transportation related issue, bigger is better, i.e. minimum mapping unit can be set to very small, even for a large coverage area. This indicates that data in GIS should have as much detailed level input as possible. Based on accuracy and precision required for specific project and available resources of time and finance, the GIS analyst should make trade off for selection of source map used to create the database.

Map scale is the ratio of map distance to ground distance (Miller and Shaw, 2001, Martin, 1996). Map scales are divided into three general categories (Miller and Shaw 2001) - Large, medium and small. Large-scale maps begin at any thing less than 1:1 and continue to 1:24000. A typical large scale map is a highway project map at a scale of 1\text{ in} \text{ represents} 500\text{ ft} (1:6000). Medium scale map range from 1:24000 to 1:100000. One of the most commonly used 1:24000 scale maps is the USGS topographic quadrangle. Any thing smaller than 1:100000 are considered small scale. Table 2.1 shows geographic extent, activity and scales adopted for the same.
TABLE 2.1 GIS-T APPLICATION AND ASSOCIATED SCALES AND ACCURACY

<table>
<thead>
<tr>
<th>Geographic Extent</th>
<th>Activity</th>
<th>Map scale</th>
<th>Data Precision, m (feet)</th>
<th>Generalized Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>State</td>
<td>Statewide planning</td>
<td>1:500000</td>
<td>250 (830)</td>
<td>High</td>
</tr>
<tr>
<td>Multidistrict</td>
<td>Corridor selection</td>
<td>1:500000</td>
<td>250 (830); 50 (170)</td>
<td></td>
</tr>
<tr>
<td>District</td>
<td>District planning</td>
<td>1:100000</td>
<td>50 (170)</td>
<td></td>
</tr>
<tr>
<td>Metropolitan area</td>
<td>Facilities management</td>
<td>1:100000</td>
<td>50 (170); 12 (40) to 9 (30)</td>
<td></td>
</tr>
<tr>
<td>Project</td>
<td>Corridor analysis, Engineering design, Construction</td>
<td>1:24000-1:12000</td>
<td>12 (40) to 9 (30)</td>
<td>Low</td>
</tr>
</tbody>
</table>

Source: Vonderoehe et al, 1993

Figure 2.3 shows the effect of changing map scale on the amount of details. For example, "Maps for ITS application may be at scale of about 1:5000 to 1:10000 in cities (similar to those used in ‘street directories’) and at smaller scales along the major roads outside metropolitan areas (Yilin, 1997).

---

![Map Accuracy Diagram](image-url)

**FIG.2.3 EFFECTS OF SCALES ON MAP ACCURACY, (SOURCE: YILIN, 1997)**
For country, state or multidistrict level planning a spatial scale of 1:100000 to 1:500000 can be used. This high level of abstraction supports budgetary planning and analysis, programme development and evaluation and policy making at upper management level. This does not require accurate and precise location data. Data can be developed at the state or county level with a grid network that covers the entire state. This level of transportation network will be primarily useful for long range strategic planning. The 20 Year Road Development Plan prepared in the country can use the spatial data at this scale (Singh, 1999). The database can be used for rail passenger, air aviation and national network of highway planning.

The feasibility studies and other aspect of project programme delivery require detail information over considerable extent. For route analysis and congestion management analysis on alternative networks a spatial scale of 1:50000 to 1:10000 is sufficient.

2.1.2.2 MAPS AND THEIR INFLUENCE ON THE CHARACTER OF SPATIAL DATA

The map is of fundamental importance in GIS transportation activities as transportation is to overcome geographic boundaries. The mapping process for different types of map is of general nature. During this process the cartographer must (Robinson et al., 1995, Heywood et al., 2004):

- Establish the purpose, the map is to serve.
- Define the scale at which the map is to be produced.
- Select the features (spatial entities) from the real world which must be portrayed on the map.
- Choose a method for the representation of these features (points, lines and areas)
- Generalize these features for representation in the two dimensions.
- Adopt the map projection for placing these features on to a flat piece of paper
- Apply a spatial referencing system to locate these features relative to each other.
- Annotate the map with keys, legends and text to facilitate the use of the map.

From the above, the scales, projections and spatial referencing are very important in GIS.
2.1.2.3 GEOGRAPHIC COORDINATE SYSTEM OF EARTH

In order to locate places on Earth, a three dimensional coordinate reference system has to be developed that takes into account the shape. The spherical coordinate system, in use for over 2000 years, is known as the geographic coordinate system, which makes use of a network of latitude and longitude (known as graticule) to fix the positions of points on the Earth.

A maps projection is a systematic representation of all or part of surface of a round body, especially the Earth, on a plane (Snyder, 1987, Lo and Yeung, 2005). When one represents the Earth on a flat plane some properties of the spherical Earth will be lost. The properties are (1) area (2) shape (3) distance and (4) direction. For spherical Earth all these four properties are correct. To achieve one or two of these properties in the final map for specific purpose of representation, different projections have been designed. The Table 2.2 shows common map projections, their properties, and major uses.

<table>
<thead>
<tr>
<th>Projection/Construction</th>
<th>Appearance</th>
<th>Properties</th>
<th>Major Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albers Equal Area/Conical</td>
<td><img src="image" alt="Albers Equal Area/Conical" /></td>
<td>Equal area; Conformal along standard parallels</td>
<td>Small Region and national Maps</td>
</tr>
<tr>
<td>Azimuth Equidistant/Planar</td>
<td><img src="image" alt="Azimuth Equidistant/Planar" /></td>
<td>Equidistant; true directions from map center</td>
<td>Air and sea navigation charts; equatorial and polar area large scale maps</td>
</tr>
<tr>
<td>projection</td>
<td>description</td>
<td>example</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Equidistant Conic/Conical</td>
<td>Equidistant along standard parallel and central meridian</td>
<td><img src="image" alt="Map of Equidistant Conic/Conical Projection" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Region mapping of mid latitude areas with east west extant; atlas maps for small countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambert Conformal Conic/Conical</td>
<td>Conformal; true local directions</td>
<td><img src="image" alt="Map of Lambert Conformal Conic/Conical Projection" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Navigation charts; U.S. State plane coordinate System (SPCS) for all east-west state plane zones continental U.S. maps; Canadian Maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mercator/Cylindrical</td>
<td>Conformal; true direction</td>
<td><img src="image" alt="Map of Mercator/Cylindrical Projection" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Navigation charts; conformal world maps</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polyconic/Conical</td>
<td>Equidistant along each standard parallel and central meridian</td>
<td><img src="image" alt="Map of Polyconic/Conical Projection" /></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Topographic maps; USGS 7.5-and 15 minute quadrangles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Projection</td>
<td>Compromise between properties</td>
<td>Thematic world map</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------</td>
<td>--------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Robinson/Pseudo-Cylindrical</td>
<td>Compromise between properties</td>
<td>Thematic world map</td>
<td></td>
</tr>
<tr>
<td>Sinusoidal/Pseudo Cylindrical</td>
<td>Equal area; local directions</td>
<td>World maps and continental maps</td>
<td></td>
</tr>
<tr>
<td></td>
<td>correct along central</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>meridian and equator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stereographic/Planar</td>
<td>Conformal; true directions</td>
<td>Navigation charts; polar region maps</td>
<td></td>
</tr>
<tr>
<td></td>
<td>from map center</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transverse Mercator/Cylindrical</td>
<td>Conformal; true local</td>
<td>Topographic mapping for areas with north south</td>
<td></td>
</tr>
<tr>
<td></td>
<td>directions</td>
<td>extent; U.S. state plane coordinate system (SPCS)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>for all north south</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>State Plane Zones</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Lo and Yeung, 2006 and Different Web pages

The Universal Transverse Mercator (UTM) Coordinate system is the most popular coordinate system among map users. The USGS decided to produce most of its digital products on the UTM (Moore, 1997). UTM system has been adopted by many organizations for remote sensing topographic mapping. The UTM grid zone designations for the world on equidistant cylindrical projection are shown in Fig.2.4.
India is located in UTM zone 41N to 47T. From these zones national grid system is designed by Survey of India (SOI) and adopted for mapping work.

Coordinate system determines the way coordinates of spatial features are stored in GIS database. Hence, while creating database, adopt the same coordinate system as in map sheet, with appropriate units. This will help in referring all the spatial element of GIS database to uniform coordinate system, and therefore allows for easy integration of spatial data.

2.1.2.4 SPATIAL DATA ACCURACY

The spatial data quality is identified by five dimensions (NCDCDS, 1988; Morrison, 1995): (i) Positional Accuracy (ii) Lineage (iii) Attribute Accuracy (iv) Logical Consistency, and (v) Completeness.

Positional accuracy is defined as the closeness of coordinate values in the geographic database to the true position of the real world features that they represent (Drummond, 1995). Conventionally, maps are accurate to roughly one line width, or 0.5mm. This is also roughly the accuracy with which an average digitizer operator can
position the cross hairs of cursor over a point or line feature. The smallest value that can be depictable or portable for an object is known as the minimum mapping unit. A 0.5 mm mapping unit or resolution is equivalent to 5m on 1:10000 scale maps. The map scales and effective resolutions are presented in Table 2.3.

TABLE 2.3 MAP SCALES AND EFFECTIVE RESOLUTIONS

<table>
<thead>
<tr>
<th>Scale</th>
<th>Effective Resolution (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:2500</td>
<td>1.25</td>
</tr>
<tr>
<td>1:10000</td>
<td>5</td>
</tr>
<tr>
<td>1:24000</td>
<td>12</td>
</tr>
<tr>
<td>1:50000</td>
<td>25</td>
</tr>
<tr>
<td>1:100000</td>
<td>50</td>
</tr>
<tr>
<td>1:250000</td>
<td>125</td>
</tr>
<tr>
<td>1:500000</td>
<td>250</td>
</tr>
<tr>
<td>1:1000000</td>
<td>500</td>
</tr>
<tr>
<td>1:10000000</td>
<td>5000</td>
</tr>
</tbody>
</table>

Source: Lo and Yeung, 2006

Weed tolerance specifies the minimum separation between coordinates while digitizing. It depends on the desired accuracy level in preserving the shape of map entities.

2.1.2.5 SPATIAL ENTITIES AND UNIVERSAL TRANSPORTATION DATA MODEL

Traditionally maps have used symbols to represent real world features. Examination of a map will reveal three basic symbol types: points, lines, and areas (Monmonier 1996, Heywood et al. 2004). These are the basic spatial entities. Each entity is a simple two dimensional model that can be used to represent a feature in the real world.

POINTS
Points are used to represent features that are too small to be represented as areas. The examples are intersections, signposts, terminals, utility poles etc. If altitude/height is important this may be represented using X, Y, Z values.

LINES
Lines are used to represent features that are linear in nature, for example, roads, rivers, pipelines, canals. The lines can also be used to represent linear features that do not exist in reality, such as administrative boundaries or international boundaries, air routes, sea navigation routes.
A line is simply an ordered set of points. It is a string of (X,Y) co-ordinates jointed together in order, and usually connected with straight lines. Lines may be isolated, such as roads, pipeline or river networks. Height or depth along line can be represented adding Z-values.

AREAS

Areas are represented by a closed set of lines and are used to define features such as fields, buildings towns, and administrative areas. Area entities are often referred to as polygons. Three dimensional areas are surfaces. These can be used to represent topographical variables such as pollutant levels, population density. Representations of entities are generalized based on the map scales. Based on this spatial entity concept, a robust data model to represent the complex relationships among the components of transportation systems is represented as shown below in Fig.2.5.

FIG 2.5 THE BASIC NODE LINK TRANSPORTATION DATA MODEL
The spatial entities may be static or mobile. Animals, vehicles and people move, and therefore, any spatial reference they are tagged with will only represent their known location at a particular time.

The spatial entities may change. River meanders, road can be relocated and policy areas redefined.

2.2 NON-SPATIAL DATA (ATTRIBUTE DATA)

Non-spatial data are attributes, associated with spatial entity point, line and area. Attributes are characteristics of an entity (Laurini and Thompson, 1992). For example, the attribute data, qualitative or quantitative, associated with spatial entities used in transportation GIS might tell us that

- A point represents intersection/bus depot/ railway station/ port
- A line represents highway/ railway track/ pipeline etc
- An area represents traffic analysis zone/ study region/ airport

Each spatial transportation entity may have more than one attribute associated with it. A point representing intersection may have attributes like number of approaches, signalized rotary traffic volume etc.

The character of attribute data influence the utility of data set in GIS analysis. The scale of measurement is one such characteristic used to record and report the data. The Table 2.4 describes the scale of measurement and the use of data.

<table>
<thead>
<tr>
<th>Data</th>
<th>Unit of measurement</th>
<th>Scale</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highway name</td>
<td>Text</td>
<td>Nominal</td>
<td>To establish identity of highway</td>
</tr>
<tr>
<td>Airport ranking</td>
<td>Value</td>
<td>Ordinal</td>
<td>To establish order or rank</td>
</tr>
<tr>
<td>Average winter Temperature</td>
<td>°C</td>
<td>Interval</td>
<td>To indicate difference but does not have origin</td>
</tr>
<tr>
<td>Size of Transportation Analysis Zone (TAZ), Area</td>
<td>m²</td>
<td>Ratio</td>
<td>Used to process population density etc.</td>
</tr>
</tbody>
</table>

Source: Heywood et al, 2004

Due to the capability of GIS to incorporate huge database, the transportation system attributes can be easily added in to database list. Transportation
system attributes are wide and large and keeps on adding. However the transportation system attributes can be classified into six groups:

(1) Physical
(2) Traffic
(3) Travel
(4) Freight
(5) Operations and Maintenance
(6) Financial

For the road transport these categories can be further divided as shown in Table 2.5

**TABLE 2.5 ATTRIBUTES OF ROAD TRANSPORT SYSTEM**

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Attribute</th>
<th>Spatial database components</th>
</tr>
</thead>
</table>
| 1      | ROAD INVENTORY  
Road geometry  
Number of lanes  
Intersection  
Type of pavement and material used  
Riding quality  
Cross drainage structure  
Shoulder type and width  
Median or central value | Link  
Link  
Node  
Node/Link  
Link  
Node  
Node/Link  
Link |
| 2      | TRAFFIC  
Traffic volume  
Traffic composition  
Average speed  
Number of accidents | Node/Link  
Node/Link  
Node/Link  
Node |
| 3      | TRAVEL  
Trip production and attraction  
Mode of travel  
Routes  
Socio Economic and demographic factors  
Trip rate/ HH | Node/Area(Zone)  
Two Nodes  
Set of Links  
Node/Area(Zone)  
Node/Area(Zone) |
| 4      | OPERATION  
Traffic sign and signal  
Road marking  
Road maintenance  
CD work maintenance  
Riding quality of the road | Node/Set of Nodes  
Link  
Link  
Node  
Link |
| 5      | FINANCIAL  
Road construction cost  
Cross Drainage construction cost  
Maintenance cost | Link  
Node  
Link |
The list of attributes is not comprehensive yet it describes the database which should be maintained by transportation corporate. The attributes are node based or linked based. They can be associated with road network using Linear Referencing System (LRS).

2.3 METADATA

Metadata is data about data. Digital data sets from data suppliers range from small scale, country and continent wide coverage available for nothing on the internet to expensive, tailor products geared to specialist market sectors (Burrough and Masser 1998). While using this existing datasets, the users need to ensure that data semantics corresponds to their uses. Thus data sets need supporting information known as metadata, about the way the survey was conducted and the information collected. A more practical level of user will consider the following information.

- The concurrency of data.
- The length of data.
- The geo-referencing system used
- The data collection technique and sampling strategy used.
- The quality of data collected.
- The data collection and interpolation methods used.
- The size and shape of the individual mapping units.
- The data exchange format.

A map legend is a pure metadata. The map legend contains information about the publisher of the map, spatial references, the map scale and its accuracy among other things. Metadata helps people to find the data they need and determine best to use it.

2.4 DATA SOURCES

National Mapping Agencies (NMA), Natural Resource Survey Institutes, Commercial Organizations, Public Works Department, and Individual Researchers are all involved in collecting, disseminating geographical data, both in analogue and digital form. Government agencies are main collector, provider and users of geographical information. Of data they collect, some 60-80 percent may be classified as geographical. The data are available as hardcopy maps, scanned images, tabular data, satellite images by remote sensing technology and Light Detection And Ranging (LIDAR), and digital
2.5 TEMPORAL NATURE OF DATA

Temporal nature of data defines concurrency and data collection cycles. Age of data decides whether the data is useful or useless for a particular end application. The environmental impact of traffic in urban areas requires current air pollution data, whereas accident trend analysis requires time series data.

2.6 DATABASE MODELLING

Data item is called an attribute or stored field. The value of attribute can be a number, a character string, a date, or logical expression i.e. 'TRUE' or 'FALSE' represented by 'T' or 'F'. Related data items are recorded and grouping of related records becomes files. The computer based data processing is based on database rather than data files. A database is an automated, formally defined, and centrally controlled collection of persistent data, used and shared by different users in an organization (Date, 1995; Elmasri and Navathe, 1994). In relational databases, data are organized in files with data items logically arranged in rows and columns, such files are called tables or relations. The coordinate values are the equivalents of data items that serve as the most elementary unit of data organization. The coordinate that forms the three basic graphical elements i.e. points, lines, polygons or areas, constitute individual records.

The real world problem of transportation is descretised into spatial objects and phenomena. The spatial objects are study area, administrative zone, traffic analysis zone, stops, road networks, rail networks, etc. The attributes of each is logically arranged by logical models like: relational data models, network data models, hierarchical data model and object-oriented models.

2.7 CREATION OF GIS DATABASE

Transportation is moving from data poor to a data rich environment. The growth of sensing, communication and information processing technologies is creating an explosion of data and many domains, including geographic and transportation data. ITS development and deployment is creating the ability to capture transportation data in real time or near real time across scale from individual vehicle to system wide flows. The
information system for transportation planning and study at regional level requires the database to be developed in a structured fashion suitable to the objectives of project. In the following sections the database development procedure for study area has been presented.

2.7.1 THE STUDY REGION

The initial study region planned is the State of Gujarat having 25 districts and 189 Talukas. The socio economic characteristics data required, transportation network data collection, digitization and bringing into required format is huge task. Considering the time and resource available, it was decided to create sample database of Anand District and the methodology is applicable for entire state for future database development.

2.7.2 OBJECTIVE OF DATABASE DEVELOPMENT

The objective of developing GIS database was to plan road users movement by different modes on the transportation network of study region and generate different scenarios. This will help to decide priority in road link improvement. Large amount of information is required for planning of regional transportation. Secondary data and primary survey using structured questionnaire are collected and structured as per the TransCAD software requirements. The database can be used to study socio-economic and trip rate study of the region.

2.7.3 SOURCES OF DATA

Transportation is multi-user multi-organized activity. The data items are not available from any single organization but from different location and different organization. The flows of data available from sources are presented in Fig.2.6. The figure shows all the possible data sources identified. However, for present study remotely sensed data are not used. The base data, both non-spatial and spatial, were obtained from the following organizations.

i) Bureau of economics and statistics, Govt. of Gujarat

ii) Census reports, Gujarat center, Govt. of India (www.censusindia.net)

iii) District Panchayat, Anand
iv) Public Works Department, Govt. of Gujarat  
v) State Police Department, Govt. of Gujarat  
vi) MORTH, Govt. of India

2.7.4 SPATIAL DATA BASE

For creation of spatial database for transportation in the study region, the jurisdiction of Anand district and Talukas are taken as polygons, highways, rails are as line and stations, bridges, intersection etc are as point entity.

2.7.4.1 MODULAR STRUCTURE OF SPATIAL DATA BASE

Transportation planning is carried out at national, state and district levels. In the State Public Work Department, the lowest level of planning is upto district and detailing of network connectivity is at village level. Hence GIS-T database created includes upto Taluka level and combining this database, the district level and state level database is created. The structure of database and development process is depicted in Fig.2.6.
2.7.4.2 SCALE

For country and state level areas, small scale spatial data of 1:1000 000 / 1:500 000 is preferred. For District and Taluka level planning large scale spatial data 1:250 000/1:50 000 are used for a detailed and better representation of features. The scale 1:250 000 is used for district level spatial database and 1:50 000 for Taluka level spatial database. The district level road-rail network is created using maps available from PWD offices, SOI map and other legacy database.

2.7.4.3 ELEMENTS OF GIS-T SPATIAL DATABASE

The main spatial elements of the GIS-T database created consist of:

- Administrative boundaries of State of Gujarat
- District boundaries
- National Highway Network in Gujarat
- Transportation zones of State of Gujarat
- Rail network in Gujarat and Anand district

For the detailed study of Anand region the spatial database created includes:

- Administrative boundaries of district of Anand
- Taluka boundaries
- Road network of Anand District including NE, NH, SH, MDR, ODR, VR
- Rail network in Anand district
- GSRTC-Bus routes for Anand Depot

The lowest level considered in the development of GIS-T database, is Taluka level and District Panchayat Electoral zone. When combined used for district level planning.

2.7.4.4 SCANNING, REGISTERING AND DIGITIZING

The various maps available are scanned, registered and digitized using Auto CAD/TransCAD software. The GPS observations at Ground Control Point (GCP) and latitude-longitude graticule on maps are used to achieve accuracy in registration. The necessary editing, if any, of geographic files of point, line and area features were done using functionalities available in TransCAD. The geographic files of the road network were then used to create the Route System which forms the Linear Referencing System (LRS)
for linking the attribute and spatial data of the road network. The different features, like Taluka polygons, point layers for urban centers, village etc were also prepared.

2.8 ATTRIBUTE DATABASE / NON SPATIAL DATA STRUCTURE

Attribute data are non spatial data associated with point, line and area entities. Attributes are characteristics of transportation zone, road-rail network, and intersection of the junction / stops.

The non spatial or attribute data were collected from all possible sources along with spatial data. The attribute was converted in to computer compatible format which can be used in TransCAD. The attribute data was structured in line with spatial database linking to generate intelligent maps. The structured database for road network prepared for study area is given in Table 2.6. Fig.2.7 shows data structure model.

![Fig.2.7 Conceptual Model for Attribute Database for a Region](image)

The attribute data associated with the road network include route names or number, capacity measures, road types, traffic volume and speed. The data for Taluka
regions includes population, socio-economic data, HH and trips per capita in regions and trips/HH.

2.8.1 ELEMENTS OF ATTRIBUTE DATABASE

The attribute data collected from various sources which are useful for deciding safety and efficiency of passenger movement, in regional context includes:

(i) The A, B, and C charts of PWD Divisions. These data includes the geometrics, pavement structures, and cross-drainage works data of road. The masonry report and inventory gives specifications of cross-drainage structures.

(ii) Form No.F1, Form No.F2 and Form.No.MTS/DS/5-0 used to compile data of MDR, ODR, VRs and traffic census there on.

(iii) Road inventory and pavement databases consist of data related to road category, administrative boundary, type of pavement surface, thickness and material used in different layers, riding quality, carriageway width, shoulder type, shoulder width, horizontal curve information, CD network information etc.

(iv) Traffic data collected by PWD twice in a year (in the month of October and May) for seven days on NH and three days on state highway, is available in time series, about 5 years.

(v) NH and SH accident record are obtained from the state Police Department. The consolidated time series accident data and spatial distribution is also recorded.

(vi) District population and population of urban centers was obtained from Census Report of 1961, 1971, 1981, 1991, and 2001. The population and socio-economic characteristics of Anand District were collected from census report at district level and other legacy databases.

(vii) Socio-economic, travel data, vehicle operation characteristics were obtained using structured questionnaires (see Annexure-I) and trained interviewers taking randomly selected samples, which are spatially well distributed.

(viii) GSRTC bus route attribute data collected from GSRTC, Nadiad Division and correlated with road network.
The attribute data, available in traffic census results, are traffic volume, traffic compositions, number of lanes, surface types have been incorporated into data base. Traffic composition and number of lanes have impact on speed of traffic and accidents. It also considers the volume of passengers and traffic on that link. The available data were structured and formatted compatible for use in TransCAD. The road network attributes data are given in Table 2.6

<table>
<thead>
<tr>
<th>ID</th>
<th>LENGTH</th>
<th>ROAD SEGMENT</th>
<th>CARS</th>
<th>JEEPS</th>
<th>BUSES</th>
<th>TRUCKS</th>
<th>TWO WHEELER</th>
<th>ANIMAL CART</th>
<th>ADT</th>
<th>CARRIAGE WAY</th>
<th>NO OF LANE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.21</td>
<td>Vadodara-Vasad</td>
<td>8204</td>
<td>6993</td>
<td>6632</td>
<td>5522</td>
<td>7878</td>
<td>94918</td>
<td>17.00</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>10.70</td>
<td>Anand-Chikhodara-Vas</td>
<td>2799</td>
<td>2787</td>
<td>24585</td>
<td>2078</td>
<td>4812</td>
<td>37061</td>
<td>12.00</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>19.92</td>
<td>Nadiad-Boravi-Anand</td>
<td>2183</td>
<td>2700</td>
<td>22938</td>
<td>1558</td>
<td>1692</td>
<td>31071</td>
<td>12.00</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>12.24</td>
<td>Kheda-Dhabhan</td>
<td>3571</td>
<td>5062</td>
<td>28197</td>
<td>1674</td>
<td>1080</td>
<td>39594</td>
<td>12.00</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>9.94</td>
<td>Tarapur-Kanawada</td>
<td>790</td>
<td>102</td>
<td>3525</td>
<td>635</td>
<td>2952</td>
<td>8004</td>
<td>3.66</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>20.19</td>
<td>Kambhat-Golana</td>
<td>691</td>
<td>219</td>
<td>2637</td>
<td>399</td>
<td>2832</td>
<td>6778</td>
<td>3.66</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>12.92</td>
<td>Gamdhina-Anklav</td>
<td>1379</td>
<td>1320</td>
<td>4188</td>
<td>1268</td>
<td>1530</td>
<td>9685</td>
<td>3.66</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>7.16</td>
<td>Vadal-Jol</td>
<td>1492</td>
<td>3102</td>
<td>4941</td>
<td>6646</td>
<td>2610</td>
<td>18792</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>23.87</td>
<td>Dhamrai-Tarapur</td>
<td>703</td>
<td>1302</td>
<td>9912</td>
<td>275</td>
<td>750</td>
<td>12942</td>
<td>12.00</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>14.16</td>
<td>Vasad-Borad</td>
<td>7432</td>
<td>4092</td>
<td>7452</td>
<td>2057</td>
<td>864</td>
<td>21897</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>22.27</td>
<td>Anand-Sojitra</td>
<td>2865</td>
<td>5640</td>
<td>7095</td>
<td>1658</td>
<td>270</td>
<td>17528</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>10.87</td>
<td>Sojitra-Limbati</td>
<td>337</td>
<td>117</td>
<td>249</td>
<td>329</td>
<td>780</td>
<td>1826</td>
<td>5.50</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>21.78</td>
<td>Nadiad-Petlad</td>
<td>1515</td>
<td>1392</td>
<td>3372</td>
<td>1136</td>
<td>804</td>
<td>8219</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>23.55</td>
<td>Petlad-Kambhat</td>
<td>1910</td>
<td>597</td>
<td>5721</td>
<td>1368</td>
<td>2346</td>
<td>11942</td>
<td>6.10</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>19.83</td>
<td>Tarapur-Kambhat</td>
<td>764</td>
<td>966</td>
<td>4482</td>
<td>435</td>
<td>3786</td>
<td>10433</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>17.12</td>
<td>Tarapur-Matankui</td>
<td>1277</td>
<td>3939</td>
<td>6420</td>
<td>911</td>
<td>3630</td>
<td>16177</td>
<td>12.00</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>14.59</td>
<td>Kambhat-Dhuvaran</td>
<td>1599</td>
<td>657</td>
<td>4791</td>
<td>1196</td>
<td>2178</td>
<td>10421</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>10.37</td>
<td>Limbati-Tarapur</td>
<td>2052</td>
<td>1149</td>
<td>3450</td>
<td>780</td>
<td>1782</td>
<td>9204</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>12.93</td>
<td>Savli-odd-Bhalaj</td>
<td>1176</td>
<td>1086</td>
<td>2355</td>
<td>587</td>
<td>798</td>
<td>6002</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>11.24</td>
<td>Uteresara-Bhalaj</td>
<td>845</td>
<td>1521</td>
<td>2665</td>
<td>530</td>
<td>2658</td>
<td>8209</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>18.20</td>
<td>Borad-Dali</td>
<td>2534</td>
<td>3837</td>
<td>8550</td>
<td>1952</td>
<td>1164</td>
<td>18037</td>
<td>10.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>8.10</td>
<td>Boravi-Anand</td>
<td>1618</td>
<td>2682</td>
<td>7407</td>
<td>944</td>
<td>312</td>
<td>12963</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>13.03</td>
<td>Anand-Borad</td>
<td>1662</td>
<td>2322</td>
<td>7698</td>
<td>746</td>
<td>318</td>
<td>12746</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4.34</td>
<td>Galiyas-Golana</td>
<td>691</td>
<td>219</td>
<td>2637</td>
<td>399</td>
<td>2832</td>
<td>6778</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>8.57</td>
<td>Borad-Bhusum</td>
<td>1659</td>
<td>2556</td>
<td>3759</td>
<td>1218</td>
<td>600</td>
<td>9792</td>
<td>9.50</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>15.89</td>
<td>Vasad-Sarae</td>
<td>1933</td>
<td>2323</td>
<td>6864</td>
<td>868</td>
<td>206</td>
<td>12323</td>
<td>7.00</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>26.97</td>
<td>Nadiad-Dakor</td>
<td>1242</td>
<td>1710</td>
<td>8211</td>
<td>985</td>
<td>858</td>
<td>13006</td>
<td>6.10</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>18.14</td>
<td>Navili-Umeti</td>
<td>1445</td>
<td>1296</td>
<td>4221</td>
<td>1333</td>
<td>1524</td>
<td>9619</td>
<td>3.66</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>3.91</td>
<td>Kheda-Vadali</td>
<td>1771</td>
<td>2232</td>
<td>7275</td>
<td>812</td>
<td>288</td>
<td>12318</td>
<td>9.50</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
2.8.3 DEMOGRAPHIC AND SOCIO-ECONOMIC DATA

The population and socio-economic data were collected from Bureau of Statistics, Anand District and census reports 1971, 1981, 1991, and 2001. The data base is used to view present scenario and generate future population based traffic and passenger movement scenarios.

2.9 GIS-T DATABASE AND UPDATING

The spatial and attribute data base created were joined using TransCAD software to create GIS-T database. All categories of road network are joined with corresponding database, which created thematic layer of road network. This can be used to view/query the different conditions of road network with name of various road segments, traffic volume, etc. Fig 2.8 shows the road network and volume composition. The attributes associated with point features (like stops, junctions, intersections) and polygons (like, TAZ, regions, District) can also be viewed in Trans CAD.

FIG. 2.8 VOLUME COMPOSITION ON SELECTED ROAD NETWORK OF ANAND DISTRICT
The database prepared requires updating periodically. The following database updating cycles can be adopted considering the data availability and resources.

**TABLE 2.7 DATABASE UPDATING CYCLES**

<table>
<thead>
<tr>
<th>Type</th>
<th>Item</th>
<th>Data source</th>
<th>Updating period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attribute Database</td>
<td>Inventory</td>
<td>Division and sub-divisions of PWD</td>
<td>As and when updated by PWD / new inventory created.</td>
</tr>
<tr>
<td></td>
<td>Traffic Characteristics</td>
<td>PWD, Bureau of Statistics</td>
<td>Twice in a year</td>
</tr>
<tr>
<td></td>
<td>Accidents</td>
<td>Police Department: Traffic Police</td>
<td>Yearly</td>
</tr>
<tr>
<td>Socio-economic</td>
<td></td>
<td>Bureau of Economics and Statistics</td>
<td>Yearly</td>
</tr>
<tr>
<td></td>
<td>Demographic</td>
<td>Census Reports</td>
<td>10 Years</td>
</tr>
<tr>
<td>Bus operation</td>
<td></td>
<td>GSRTC Schedules</td>
<td>As per up gradation of schedules</td>
</tr>
<tr>
<td>Spatial Database</td>
<td>Road/Rail network</td>
<td>PWD, SOI</td>
<td>20 Years</td>
</tr>
<tr>
<td></td>
<td>Administrative Boundaries,</td>
<td>SOI</td>
<td>As updated by SOI</td>
</tr>
<tr>
<td></td>
<td>Topography</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Villages and other</td>
<td>SOI and Census reports</td>
<td>Every 10 years as per census reports As per changes in routings</td>
</tr>
<tr>
<td></td>
<td>Settlement locations</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GSRTC Bus routes</td>
<td>GSRTC Divisions</td>
<td></td>
</tr>
</tbody>
</table>

**2.10 GIS-T IMPLEMENTATION: HUMAN AND ORGANIZATIONAL ISSUES**

The stack holders involved in movements of road users are the road network users (GSRTC, Private operators, Individuals), network providers (PWD and BOT Participants), Law enforcement agencies (Police Departments, Air Pollution Control Agencies). The GIS-T must serve the needs of these stack holders and seamlessly and effectively fit into information technology and Decision Support System (DSS) cultures of
these organizations. The organizations are maintaining their databases with the help of existing man power and structure of database prevailing in tabular formats and hard copy files. The stand alone PCs are used to maintain the attribute database and compared with hard/soft copy maps for spatial information. The existing information technology infrastructure of an organization will affect the nature of any GIS implementation. The other factors are cost of software, data creation, range of functionality and need of organization, and availability of man power and training of in house manpower and support systems.

In the country, Public Works Department is the main organization responsible for planning and management of 3.3 million kilometer length of road network in coordinated transportation planning with Railways and Air Aviation facilities.

Thus both PWD and Build Operate and Transfer (BOT) participants to maintain the categorised road network in efficient and operational conditions and deciding priority database is important. For passenger and freight movement operators, road network with good riding quality, better geometrics with safe and desired speed is to be available for optimization of minimum path and high return and revenue generation. For enforcement agencies spatio-temporal data of events are necessary.

To meet the data requirements the filed data are used, which becomes laborious, time consuming, unreliable and even not available in time to decision makers. For complex activities the tabular data format becomes bulky and standard size of sheets and printers are limitations for presentation of information for monitoring and operation purposes.

The GIS database developed in the present study will be immensely useful to the user organizations. The integration and flow of data/data sharing for overall intelligent transportation system (ITS) is proposed as shown in Fig. 2.9.
FIG. 2.9 FLOW OF DATA AND DATA SHARING IN ITS
2.11 GIS-T DATABASE DEVELOPED

It is evident from the above discussion that the GIS-T database development is a challenging task due to huge data collection, processing to computer and software compatible format and various organizational impediments during the process. The detailed road network data available with the sub-divisional level is difficult to collect due to policy framework and decision making at lower level. The same problem is faced during data collection from GSRTC- the route network details- and police department -the accident records. The socio-economic and travel data collected by primary survey is huge in nature and good amount of resources were required. The database created can be used by several stakeholders in transportation and addition of data can be feasible in present database.

2.12 CONCLUSIONS

The GIS-T database developed needs huge amount of information, which are spread over various organizations. To collect the detailed data large resources are required, but the investment is worth as the same database can be used by several organizations.

The methodology described can be used for creation of similar databases. The spatial scale used for present GIS-T database is 1:250000 which is suitable for regional road network planning, GSRTC route network planning and district level road accident management system development. The database created with spatial scale of 1:50000 is useful for Taluka level planning, ODR, VR management and village level connectivity planning.

Whenever the additional information is available the same can be added to the database. As the government is heading towards e-Governance, the database will be immensely useful to decision makers at all level of administration. GSRTC can plan their regional transit bus routes based on shortest paths for point to point service and local service based on most economical routes and passenger demand.