

UC Santa Barbara

GIS Core Curriculum for Technical Programs (1997-1999)

Title

Unit 3: Locating Transportation Data

Permalink

<https://escholarship.org/uc/item/5fj0v32d>

Authors

3, CCTP
Noronha, Val

Publication Date

1998

Peer reviewed

UNIT 3: LOCATING TRANSPORTATION DATA

Written by Val Noronha, Digital Geographic Research Corporation,
Mississauga Ontario, Canada

Context

Transportation data are used in a widening variety of applications, e.g. to "spot" customers in marketing studies, track emergency vehicles and to optimize delivery routes. Recently it has become possible to monitor traffic volumes and speeds on highways, and to pass advisory messages to dashboard computers in vehicles. This emerging technology is Intelligent Transportation Systems (ITS), formerly Intelligent Vehicle/Highway Systems (IVHS).

It is useful to think of **five** components of transportation data:

- a. the coordinates of the street centerline (a polyline);
- b. the name or label of the centerline, e.g. Erin Mills Parkway
- c. address ranges on either side of the road;
- d. topological connectivities - whether adjacent links are accessible
- e. attributes (e.g. pavement quality, speed limit, traffic volume);

data on landmarks and facilities such as restaurants are increasingly stored as attributes in conjunction with street network data). With these data, one can:

- a. construct a map on screen or hardcopy from a digital database;
- b. using the street name, orient oneself on the map;
- c. geocode an address, i.e. determine the coordinates of 421 Erin Mills Parkway;
- d. determine an optimal route - shortest, quickest, scenic, easiest
- e. perform advanced analyses, e.g. schedule road maintenance, optimize the use of a vehicle fleet to pick up and deliver goods, model the consequences of closing a major traffic artery at rush hour, etc.

The availability and quality of data are critical to the success of an application. When the required data are not available, surrogate data are often used, e.g. one might use drive distance rather than drive time to compute the quickest route between points. It is important to understand the limitations of the analyses when data are substituted in this way.

To acquire data appropriate to a transportation application, the student must first understand the different types of data used, then survey the market critically to find the product most appropriate to his needs.

Example Application

A medical emergency is called in to a 911 center. Emergency personnel need ...

- **to know the precise location of the incident;**
- **to determine the available vehicle that could respond most quickly, and the treatment facility that could be reached most quickly from the incident location;**
- **to recommend routes to the ambulance driver if required, and to determine alternate routes in the event of congestion or some other unanticipated situation.**

There are two components to the problem:

1. **to geocode the incident location**
2. **to compute shortest routes**

Accordingly, two kinds of data are required: (1) street names, address ranges and landmarks, and (2) drive-distance or drive time attributes, and topology. The example is analogous to home delivery for a department store chain or pizza outlet, or a car pooling problem.

The GIS technician will identify and evaluate possible sources of data, taking into consideration origin and quality of the data, and impact on results. He will perform the address matching and route optimization operations as well, but the technical details of these are beyond the scope of this unit.

Learning Outcomes

The following list describes the expected skills which students should master for each level of training, i.e. Awareness/Competency/Mastery.

Awareness:

To learn about potential applications of transportation data, and the appropriate data types and typical sources for each class of application.

Competency:

To learn about acquisition and processing of data, by means of a practical exercise.

Mastery:

To assess data quality, and to anticipate limitations on accuracy of output; to appreciate cost and benefit issues.

Preparatory Units

Recommended:

1. Unit 7 - Using and interpreting metadata
2. Unit 8 - Error checking
3. Unit 9 - Converting digital spatial data between formats, systems and software

Complementary:

1. Unit 46 - Address matching
-

Awareness

Learning Objectives:

1. Student **understands** essential concepts and vocabulary of street networks.
2. Student broadly **understands** how basic network operations (map construction, address matching, shortest path calculation) are performed by commercial GIS, what data are used and how they are processed.
3. Student can differentiate between data types required to perform the various operations in (2).
4. Student can list most common sources of data, and characteristics of each source, with particular emphasis on method of data gathering, and quality.

Vocabulary

- centerline
- address range
- address match
- geocode
- topology
- attribute
- impedance
- shortest-path
- planar
- turn table
- GPS

Concepts

- geometry and structure of polyline, and related GIS issues (coordinate transformations, length measurement, clipping and edge matching at tile boundaries)
- standard street name components (e.g. in 4215 Erin Mills Parkway North)
 - proper name: Erin Mills
 - street type: Parkway
 - direction prefix/suffix: North

- addressing
 - addressing styles: 4215, 4215A, 4215½
 - odd and even addresses, and what they mean; variability in conventions
 - block-based and distance-based address assignment
 - address points
 - address ranges
 - rural addressing
- address matching methods
- impedance measures
 - distance
 - difference between driven distance and digitized distance - most severe on sinuous mountain roads because of cartographic generalization of the ucurve
 - legal issues: one ways and turn restrictions, HOV lanes; difficulty in maintaining currency
 - transit time: delays at intersections due to signals; variability with time of day/year
- database implementation methods
 - turn tables
 - non-planar intersections: data storage using turn tables
- incident reporting, video cameras and inductive loop detector technology
- driver instructions: radio reports, overhead advisories, dashboard displays and multimedia

Data Sources

Street network data are available in varying degrees of quality and completeness.

- **Centerlines** are surveyed by federal government agencies, at about 1:100,000. These files are typically produced under good quality control, but have no attribute information (e.g. street names and address ranges) attached, and topology may or may not be captured. For larger scale data one would typically have to approach state and municipal governments.
- **Multipurpose Street Network** files (i.e. street name, address ranges and topology stored) may also be produced by the federal government (TIGER U.S. Bureau of the Census; Street Network Files Statistics Canada). These databases are not always suited to specific applications. They were originally developed to assist census taking efforts; their structure and content reflect this limited utility.
- **Attribute Sets** such as speed limit and travel time are expensive to gather and to maintain, and hence difficult to acquire. Private sector data vendors in some countries offer value-added products, usually based on government multipurpose street network files. There is a wide variety in the quality of these products. For example, speed limit, number of lanes and travel time are sometimes inferred based on the type of neighborhood. Since the early 1990s, real-time highway flow data (volume and velocity) have been gathered in many cities using inductive loop detectors embedded in the pavement; these data may be available from state-controlled Traffic Management Centers (TMCs) or private agencies.

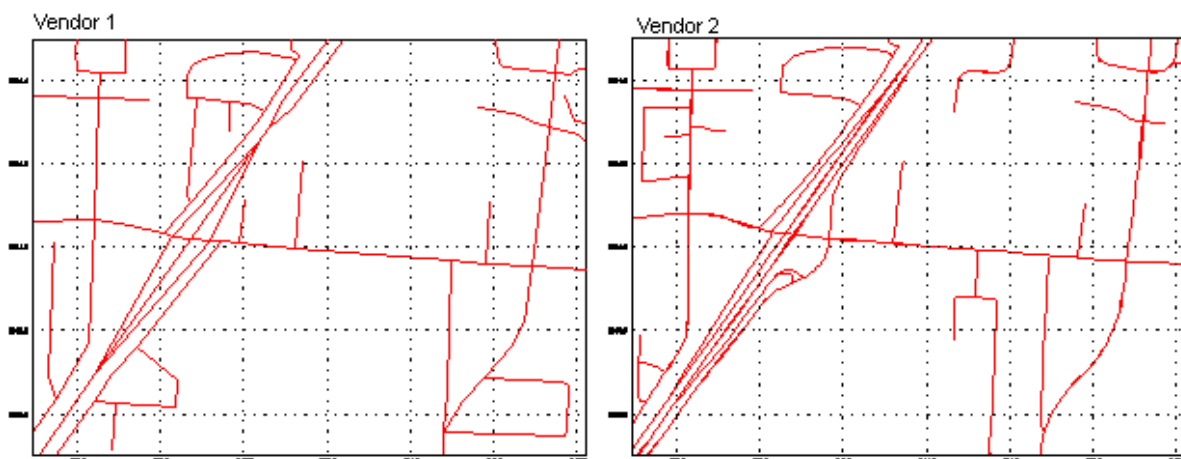
The following table is necessarily skeletal and incomplete. The inclusion of a vendor name does not constitute endorsement of data quality or the firm's reputation; similarly, omission of a name should not be construed as disapproval of a firm or its product. Resellers of unmodified government data are not cited, neither are those who exclusively offer data packaged with software for the mass market (e.g. Rand McNally's Street Atlas USA).

	Government	
--	-------------------	--

Country	Government (centerline)	(Multipurpose Street Network)	Private
Australia	<ul style="list-style-type: none"> • AUSLIG • State mapping agencies • Municipalities 	<ul style="list-style-type: none"> • Victoria: Land Victoria 	<ul style="list-style-type: none"> • P-Data? (Sydney) • UBD (Sydney) • Melway (Melbourne)
Canada	<ul style="list-style-type: none"> • Natural Resources Canada • Provincial mapping agencies • Municipalities 	<ul style="list-style-type: none"> • Street Network File (Statistics Canada) 	<ul style="list-style-type: none"> • Compusearch (Toronto) • Desktop Mapping Technologies (Markham ON)
USA	<ul style="list-style-type: none"> • DLG (USGS) • National Highway Planning Network • State mapping agencies • Municipalities 	<ul style="list-style-type: none"> • TIGER (US Census) 	<ul style="list-style-type: none"> • Etak (Menlo Park CA) • GDT (Lebanon NH) • Navigation Technologies (Sunnyvale CA) • Thomas Brothers (Irvine CA)

Example Implementation

Figure 1 shows a typical centerline street network for a portion of Santa Barbara, California, USA, as represented by two commercial data vendors. There are obvious visual differences. Clearly Vendor 2 has taken far greater trouble to represent the dual-carriageway freeway and the exact geometry of the exit ramps; and each vendor has included roads that the other has not (usually driveways in semi-private buildings). However and this is not evident from the illustration database 2 is not topologically structured, and street names are stored in a haphazard format, e.g. the main east-west artery is labelled Hollister Av in some places, and simply holister in others. Such data, although positionally accurate, cannot be used for geocoding or routing.



Also not apparent from the illustration is the fact that Vendor 1 shows intersections between the

freeway and the arterial road, although in reality the freeway overpasses the artery. This is because the vendor's data structure cannot distinguish between planar (at-grade) and non-planar intersections. Some vendors may use turn tables to store such data.

In sum, although Vendor 1 leaves much to be desired in positional terms, the data can be used for topological applications such as address matching and routing. Vendor 2 is good only for graphic applications such as map display.

Competency

Learning Objectives:

1. Student compiles detailed list of data requirements, including quality requirements, for a given application.
2. Student surveys market for data sources, and selects a vendor.
3. Student prepares data in format required for the application.

Vocabulary

- Conflation

Concepts

- Typical formats for data exchange
 - ArcInfo Export (E00)
 - ArcView shape (SHP/SHX/DBF)
 - DLG
 - MapInfo (MID/MIF)
- Typical coordinate systems
 - UTM
 - Latitude/Longitude
 - Lambert
 - State Plane (USA only)
- Organization: Roads are typically organized by class, using a feature code field in the attribute table. There are wide differences between vendors.
 - Some include non-road layers such as hydrography and landmarks, others restrict themselves to transportation layers.
 - Some classify roads as freeways, major arteries, collectors, residential streets, etc. There are differences in taxonomy, and it is virtually impossible to produce identical subsets of streets in any class, from different vendors.
- Synthesis: It is not always possible to find all the data required for an application, at a single source. One may have to blend the strengths of one database with those of another: e.g. linework from municipal sources, attaching attributes from a private vendor; or older neighbourhoods from one vendor and newer subdivisions from another. This is conflation.
- In the ITS arena, although the infrastructure of loop detectors is widespread even in medium sized cities, the data from these sensors are currently not well integrated with street networks.

Tasks

For a given application, student compiles detailed list of data requirements, particularly attribute requirements.

- Student surveys market for data sources, compiling a table of product features offered by various vendors.
- Using data available at institution, student prepares data in format required for the application.



Example implementation

To calculate drive time database: develop travel time attributes based on (a) zonal assignment of impedance values, (b) field notes and local knowledge, (c) vendor data if available, (d) TMC data if available. Feed the data into a system to perform optimal routing, and for one or two origin-destination pairs, compare the system recommendations against real drive time measured in a vehicle.

Mastery

Learning Objectives:

1. Quality issues
2. Cost and benefit issues

Vocabulary

- Accuracy
- Precision
- Resolution
- Coverage
- Currency
- Metadata

Concepts

- Data structure native to a GIS may not accommodate local municipal practices, or analytical requirements of organization:
 - Multiple names for a street (Macdonald Freeway, Hwy 401): creates difficulties in geocoding
 - Address schemes may differ on two sides of street (2000-2100 on south side, 1701-1799 on north) if municipalities are different
 - Linear referencing, dynamic segmentation.
- Better quality data can always be created, if sufficient funds are available, or by pooling funds between interested agencies. The latter solution requires that the product satisfy the requirements of all subscribing parties.

Tasks

Trace the history of street network data available for a selected area, examining the parties that created it, and their needs, compared with current and emerging needs.

Follow-up Units

Suggested:

1. **Unit 30 - Validating databases**
 2. **Unit 31 - Managing database files**
-

Resources

[outdated links have been removed]



Data, documentation, and commentary.

TIGER U.S. Census Bureau ITS Online ITS new



Examples of network-based applications.

Mapmaker for U.S. street addresses



Real time freeway data

Los Angeles, California
Toronto, Ontario, Canada
Paris, France

Created: May 14, 1997. Last updated: October 5, 1998.