NCHRP
National Cooperative Highway Research Program

RESEARCH RESULTS DIGEST

September 1997 Number 218

Subject Area: IA Planning and Administration Responsible Senior Program Officer: Kenneth S. Opiela

A Generic Data Model for Linear Referencing Systems

This NCHRP digest describes the findings of NCHRP Project 20-27(2), Systems and Applications Architecture for GIS-T, conducted by Alan Vonderohe, Chih-Lin Chou, Forest Sun, and Teresa Adams, Department of Civil and Environmental Engineering, The University of Wisconsin—Madison. The digest was prepared by Kenneth S. Opiela, Ph.D., P.E., NCHRP Senior Program Officer, from the contractor's interim report.

INTRODUCTION

This digest describes a consensus location referencing data model that was developed under NCHRP Project 20-27(2), *Systems and Applications Architecture for GIS-T*. The model allows linkages to various types of data over all modes. This information will be of use to persons involved with the design and implementation of field location referencing systems as well as the structuring of agency databases for location referencing.

State departments of transportation (DOTs) have been adopting Geographic Information Systems (GIS) over the past decade with varying degrees of intensity and success. DOTs manage vast stores of linearly referenced data that must be associated with cartographic representations to be used in a GIS. Thus, there is a fundamental need for a generic data model for Geographic Information Systems for Transportation (GIS-T) to provide the linkage to linear referencing components. Linear referencing systems are used in nearly all application areas that are based upon networks, including infrastructure management, transit, freight, intelligent transportation systems (ITS), waterway navigation, hydrological analysis, utilities management, and seismological sensing.

BACKGROUND

Previous work on NCHRP Project 20-27 led to the recommendation that transportation agencies develop conceptual organizing principles founded upon the notion of location as a data integrator (Vonderohe et al., 1993). Another result of NCHRP Project 20-27 was a suggested technological framework (server net) for support of GIS-T and transportation computing in general. Generic functional and data models to complement the technological model developed in the initial project will be developed under NCHRP Project 20-27(3).

The significance of linear referencing methods and systems to transportation applications has been recognized for some time. *NCHRP Synthesis of Highway Practice 21*, "Highway Location Reference Methods" (1974), made the distinction between methods and systems, classified a number of linear referencing methods, and made recommendations for their improvement. Transportation agencies from time to time have studied the location referencing methods they use and sought to adopt standards for them (Briggs and Chatfield, 1987). During one study, the Michigan DOT identified 38 location referencing methods in use by the agency. More recently, some state DOTs

CONTENTS

10-29-91

have developed formal data models for location referencing (Deighton and Blake, 1993; Ries, 1993; Scarponcini, 1994; Rowell, 1996). Some have succeeded in raising the issue to the policy level (Deighton and Blake, 1993). A number have developed models and procedures for including linearly referenced data in GIS (e.g., Cihon, 1996). The Wisconsin DOT has recognized "location control management" as a formal business area in its information strategy plan and developed rigorous procedures for management of linearly referenced data (WisDOT, 1996).

Early research in GIS in transportation led to identification of the need for, and subsequent development of, dynamic segmentation as a critical function for managing linearly referenced data (Fletcher, 1987; Dueker, 1987; Nyerges and Dueker, 1988; Nyerges, 1990). More recently, the underlying data models that support current implementations of dynamic segmentation have been examined (Dueker, 1992).

An executive-level commitment to the concept of a National Spatial Data Infrastructure (Mapping Science Committee, 1993) spurred interest in the development of standards and common models for data. The Ground Transportation Subcommittee of the Federal Geographic Data Committee recommended that a linear referencing system be incorporated in standards efforts. The Federal Highway Administration (FHWA) incorporated linear referencing systems from the states in the Highway Performance Monitoring System (FHWA, 1993)

At the same time, a number of workers addressed various aspects of the data conflation problem, typically associated with attempts to integrate census data tied to TIGER/line files with attribute data tied to other representations of the same street network that do not coincide with TIGER (e.g., Brace and Peterson, 1994; Clark and Bain, 1994; Peterman, 1994;).

Given all these activities and interests in data sharing and integration, the need for a common, generic data model for linear referencing systems was compelling. Thus, a workshop was convened to address this need.

WORKSHOP

On August 5 and 6, 1994, 42 transportation professionals, systems developers, and academics attended a workshop in Milwaukee, Wisconsin, with the objective of preparing a draft consensus conceptual data model, at the entity-relationship level, for linear referencing systems. Workshop participants, selected for their expertise in linear referencing systems and modeling, represented local, state, and federal transportation and mapping agencies; consultants, data providers, and software providers from the private sector; and researchers from national laboratories and universities. Recognizing that it was not feasible to develop a data model that would meet all the needs of all application areas, a generic model was sought that met common needs and formed a core that could be extended as needed in specific application areas.

The resulting draft data model, in object modeling form, associates transportation data with multiple cartographic representations and multiple network models through a single linear datum. The datum links the data model to real-world features and provides the referencing space that enables transformations among linear referencing methods, networks, and cartographic representations at various scales. The data model supports a set of fundamental operations that cause data to flow between the database world and the real world. The data model as presented is intended to represent the requirements for a linear referencing model—it is not intended as a specification.

A first-draft report was prepared from notes and other materials developed during the workshop, audio tape recordings of workshop sessions, and follow-up discussions with workshop participants. All participants were given the opportunity to review the first draft and provide responses. These responses were assimilated to identify both consensus revisions to the first draft and significant points of contention. The responders were informed of these results and asked to provide their opinions on each point of contention. The responders were also asked to provide measures of the relative importance of their positions on each point of contention (i.e., "critical," "strong preference," "weak preference"). The revised model presented in the second-draft report was true to the model as developed during the workshop and to the revisions on which there was consensus. The second-draft report was presented and published in the Proceedings of the 1995 AASHTO GIS-T Symposium in Sparks, Nevada (Vonderohe et al., 1995). Since that time, additional research, by both the authors and other, independent researchers, has led to refinements in the data model that are incorporated herein. Furthermore, the topic of location referencing in general and linear referencing in specific has attracted a great deal of interest and discussion since the initial workshop. Implementation issues and methods for measurement system design are among the recently discussed ideas. Most recently, the TRB Task Force on GIS-T opened an unmoderated World Wide Web-based discussion on the subject of linear referencing. Components of a dialog, included in that Web-based discussion, between staff at the Bureau of Transportation Statistics (BTS) and the principal author of this report, are presented in the discussion section of this digest.

Since publication of the second-draft report, other researchers have discovered the need for refinements in the data model that have been incorporated herein. These researchers include David Fletcher (Alliance for Transportation Research), who discovered a necessary refinement when implementing the model during Phase B of the GIS-T Pooled Fund Study and Todd Hepworth (University of Wisconsin—Madison), who when working with the principal author of this report (Vonderohe) on research funded by Sandia National Laboratories, independently discovered the same refinement as Fletcher.

Wende O'Neill (Utah State University) and Bruce Spear (BTS) conceived of a dialog on linear referencing that is part of a World Wide Web-based discussion. They developed the questions and adapted the responses of the principal author of this report into a Web-based discussion document, components of which have been subsequently freely incorporated in this digest. Necessary clarifications and revisions to some definitions, which became apparent during development of the dialog, have also been adopted.

Workshop Structure

All participants received a package of materials in advance of the workshop. The package included information on the research project, the objectives of the workshop, a preliminary program, and a set of pre-conference papers whose authors had been asked to give presentations.

A modified form of the "Technology of Participation" method was used to structure the workshop, with Professor Tim Nyerges of the University of Washington serving as facilitator. Following introductory remarks concerning workshop methods, 12 invited technical presentations were made concerning various aspects of linear referencing systems data modeling. During the presentations, participants identified issues and wrote them on large cards that were posted during breaks. Following the presentations, participants were encouraged to identify additional issues and post them. The issues were then clustered into topic areas by the group as a whole. The clustered issues were then synthesized and gaps were identified.

It was decided that the topic areas "Terms and Definitions" and "Scoping" were most critical to development of the model, and a discussion of these topics by the group ensued. These discussions ultimately led to collective development of a linear referencing system data model, which appeared in the first-draft report. The data model presented in this digest includes consensus revisions based on responses to the first draft provided by workshop participants. It also includes refinements identified as necessary through independent research since initial publication of the second-draft report.

Topic Areas and Issues

Critical topic areas and issues identified during the workshop included the following:

1. **Terms and Definitions.** Example issues: Standardized, unambiguous definitions must be developed for common terminology. Terms such as "traversal" should be used instead of "route," "path," or "trip," all of which might be subclasses of "traversal." Terms such as "anchor point" and "anchor section" should be used because "control point" and "control section" have other meanings. The term "distance" can mean "odometer distance," "posted distance," or "cogo distance."

2. **Scoping.** Example issues: Are geographic, spatial, cartographic, and temporal objects modeled in the same domain? What is the conceptual extent of the term "linear referencing system"? Is there a set of core requirements for a host of applications? We must account for vector, non-planar models used in commercial GIS. What are the primary functions that must be supported? Is *linear* referencing broad enough to address workshop goals?

3. **Schema Constructs.** Example issues: What are the primary building blocks of linear referencing methods? What are the differences between structural data model requirements and functional views? Are linear referencing system components hierarchical objects or interdependent and relational? We must relate topology, one-dimensional, two-dimensional, and three-dimensional space. Can one conceptual data model accommodate feature-based systems, planar graph-based systems, and non-planar graph-based systems?

4. **Transformation.** Example issues: How do we uniquely and unambiguously identify locations for transforming that information between dissimilar information systems? What are the basic transformations between linear referencing methods? How do we link multiple linear referencing methods together for data integration in the network domain? What are the rules for aggregation to support more generalized reporting?

5. **Multi-Dimensionality.** Example issues: To address conflation, (1) should we develop unique identifiers for certain features; (2) should we standardize on topology or geometry (x,y,z)? Do we need a Global Positioning System (GPS)/GIS linkage? Linear referencing systems must be linked to higher-dimensional systems, including those that model time. Spatial proximity is not a surrogate for network topology.

6. **Methods and Coding.** Example issues: How are the "lowest common denominator" sites identified? What is the appropriate datum structure? What is meant by the "location" of a bridge—is it the center, one end, the other end? The model must be able to handle very large databases. What are the rules for establishing linear referencing method starting and ending points? What is the best method for referencing ramps? What geographic features are assigned external identifiers?

7. **Data Integrity.** Example issues: How do we ensure data integrity if we have multiple linear referencing methods? What are the referential integrity rules? Can versions be coordinated by adding version numbers to unique identifiers? What are the implications of alignment changes?

8. **Institutional Policy.** Example issues: Do users need to understand linear referencing systems? What policies and procedures are required for a linear referencing system? What cost constraints are associated with a linear referencing system?

Key Concepts

NCHRP Synthesis of Highway Practice 21 contains two fundamental premises adopted by the workshop participants:

1. There is a clear distinction between *linear referencing methods* and *linear referencing systems*: "A highway location reference *system* is a set of office and field procedures that includes a highway location reference *method*. The latter is a way to identify a specific location with respect to a known point." The workshop participants included within the concept of *system* a means for transformation among various *methods*. Thus, "milepoint," "reference post," and "engineering stationing" are methods. The policies, records, and procedures that relate these methods are the system.

2. The location of any unknown point along a linear feature can be determined by specifying the direction and distance from any known point to the unknown point. All linear referencing methods are based on this. The workshop participants concluded that the premise is true for two and three dimensions also. Each additional dimension removes a constraint on direction.

It was concluded that the time limitations of the workshop precluded development of a data model that supported all the needs of all possible application areas. Therefore, a core, generic data model was sought that could be extended to meet specific needs of various applications. Additionally, a model that addressed the *requirements* for linear referencing systems was pursued rather than a robust and elegant *specification*.

Multiplicity became a theme. There is a central need to integrate not only multiple scales of geography and cartographic (coordinate-based) data from multiple sources, but also multiple network models, each of them necessary for particular applications.

It was decided that the data model must support the following fundamental operations:

1. **Locate.** Establishment of the location of an unknown point in the field by reference to objects in the "real world."

2. **Position.** Translation of a real-world location into a database location.

3. **Place.** Translation of a database location into a real-world location (the inverse of the "position" operation).

4. **Transform.** Conversion between various linear referencing methods, represented by database locations; between various cartographic representations; and between methods and cartographic representations.

It was expected that if these operations are supported, then the model should support higher-level operations such as those associated with GIS (e.g., overlay, connectivity, proximity) and those associated with network analysis (e.g., pathfinding, routing, location, and allocation).

DATA MODEL

Overview

Figure 1 presents a conceptual overview of the data model. The central notion is that of a linear datum that supports multiple cartographic representations (at any scale) and multiple network models (for various application areas). The datum provides the fundamental referencing space for transformations among various linear referencing methods, network models, and cartographic representations. It also links the model to the "real world" through attributes that describe its location and spatial characteristics in real-world references and measures.

Cartographic representations provide coordinate references, the basis for to-scale visualization of the model, and linkages to two-dimensional and three-dimensional GIS databases. Network models provide the topological framework for pathfinding, routing, location/allocation, transshipment, and flow operations.

A number of linear referencing methods might be associated with each network model. These methods might be those associated with infrastructure management, such as reference post, milepoint, or engineering stationing. They might also be those associated with navigation (requiring recognizable landmarks or navigation aids), or with transit (timing points), or with a host of other application areas. Each linear referencing method ties a collection of business data to the model, thereby providing a means for integration of those data. The linear referencing system can be thought of as all those components of the model that provide methods for location referencing of business data, transformations among those methods, and linkage of the model to the "real world" and its cartographic representations.

The object model diagram appears in Figure 2. The diagramming method is slightly modified from that of Rumbaugh et al. (1991). The modification being that low and high cardinalities are shown with Arabic numerals on both ends of all associations. Standard notation includes the name of an object class in the upper half of a rectangular box; attributes of the class in the lower half of the same box; an association between two object classes denoted by a line connecting the boxes; an association descriptor written on the connecting line for all associations except aggregations; attributes of an association appearing in the lower half of a rectangular box tied to the association's connecting line by a half loop; "many" cardinality indicated by a filled circle; "zero or one" cardinality indicated by an empty circle; "exactly one" cardinality indicated by lack of a circle; and aggregation indicated by a diamond symbol.

Object Classes and Their Attributes

Linear Datum. The complete set of anchor sections and anchor points, constituting a mutually exclusive, totally exhaustive, ordered set of linear locations[1] (see Figure 3). The linear datum relates the database representation to the real world and provides the domain for transformations among linear referencing methods and among cartographic representations. There is a single linear datum. It is

[1] Definition derived from personal communication with David Fletcher (March 1997).

included in this data model because of the centrality of its concept to the overall model, not because there would necessarily be a number of instances that would have to be tracked in a database. Various versions of the linear datum might exist over time as changes in transportation facilities occur. No attributes are assigned to the linear datum.

Anchor Point. A zero-dimensional location that can be uniquely identified in the real world in such a way that its position can be determined and recovered in the field. Each anchor point has a "location description" attribute that provides the information necessary for determining and recovering the anchor point's position in the field. Forms of location descriptions can vary and can be quantitative or descriptive or both, (e.g., the intersection of the centerlines of Oak Street and Maple Street; and 1.2 miles south of the Post Office on the centerline of Route 9).

Anchor points can be understood as one-dimensional control points, in that they serve the same purpose as geodetic control points in two and three dimensions. That is, they are the fundamental objects to which all other objects are directly or indirectly tied.

Anchor Section. A continuous, directed, nonbranching linear feature, connecting two anchor points, whose real-world length (in distance metrics), can be determined in the field. Anchor sections are directed by specifying a "from" anchor point and a "to" anchor point. Anchor sections have a "distance" attribute, which is the length of the anchor section measured on the ground. Values are expressed in units of linear distance measure (e.g., kilometers).

Anchor sections provide the fundamental referencing space. The collection of anchor sections in a given linear referencing system is analogous to the ellipsoid surface in a geodetic datum or the map projection surface in a two-dimensional Cartesian referencing system.

Cartographic Representation. A set of lines that can be mapped to a linear datum (see Figure 4). The set of lines can be either fully or partially linked. That is, the set can consist of disjoint groups with the lines in each group being internally linked. Cartographic representations have a "source" attribute that denotes the source (scale and lineage) of the object. Scale values are expressed as ratios or as equations that relate distances measured on the source form of the cartographic representation to distances measured on the ground.

Cartographic representations provide coordinate references; the basis for to-scale visualization of other components of the linear referencing system model; and linkages to extended topological, vector- based GIS data models.

Line. "A generic term for a one-dimensional object" (USGS, 1992). Spatial Data Transfer Standard (SDTS) goes on to define five specific kinds of lines: (1) line segment, (2) string, (3) arc, (4) link, and (5) chain. A line, as defined herein, can be any of these except a link. This is because lines, as defined herein, have a "shape and position" attribute. According to SDTS, a line segment is a direct line between two points, a string is a connected nonbranching sequence of line segments, an arc is a locus of points that forms a curve that is defined by a mathematical expression, and a chain is a directed nonbranching sequence of nonintersecting line segments and (or) arcs bounded by nodes, not necessarily distinct, at each end. Shape and position are provided either by the x,y,z coordinates of points associated with line segments or by the mathematical expressions associated with arcs. Possibilities for types of coordinate values include Cartesian and geographic (lat/long/elev). Possibilities for mathematical expressions include splines and polynomials.

Network. A graph without two-dimensional objects or chains. If projected onto a two-dimensional surface, a network can have either more than one node at a point and (or) intersecting links without corresponding nodes. Note: This is a modification of the definition provided by the SDTS. Modification is necessary to exclude chains. Within the context of the linear referencing system data model, a network is an aggregate of nodes and links and is, thus, a purely topological object (see Figure 5). The network component of the model provides the basis for analytical operations such as pathfinding and flow. No attributes are assigned to networks.

Node. A zero-dimensional object that is a topological junction of two or more links, or an end point of a link. Note: This is a modification of the definition provided by the SDTS. Modification is necessary to remove reference to chains. In this data model, nodes do not have coordinates. They are located geometrically by reference to the datum.

Each node has a "datum measure" attribute that is used to locate it on an anchor section. "Datum measure" is an offset measured from the "from" anchor point of the anchor section. "Datum measure" is expressed as a distance measure in the same units as the "distance" attribute of the associated anchor section.

Link. A topological connection between two ordered nodes. Note: This is a modification of the definition provided by the SDTS. Modification is necessary to require directionality. Each link has a "weight" attribute that is a linear measure of impedance associated with travel along the link. Weights are often expressed in distance measure, but they could be in other linear metrics such as travel time or cost.

Linear Referencing Method. A mechanism for finding and stating the location of an unknown point along a network by referencing it to a known point. Note: This is a modification of the definition provided by Deighton and Blake (1993). There are many kinds of linear referencing methods (e.g., milepoint, reference post, and engineering stationing). All linear referencing methods consist of traversals and associated traversal reference points that together provide a set of known points, a metric, and a direction for referencing the locations of unknown points (see Figures 6 and 7). No attributes are assigned to linear referencing methods.

Traversal. An ordered and directed, but not necessarily connected, set of whole links. Coding conventions are required for establishing traversal directionality (in contrast to link directionality) and for specifying nonconnected traversals. No attributes are assigned to traversals. Note: It was the intent of the workshop participants to allow dendritic traversals, but specific implications of this for the model have not been investigated.

Traversal Reference Point. A zero-dimensional location along a traversal that is used to reference events along the traversal. Each traversal reference point has a "traversal measure" attribute, which is used to locate it along the traversal. "Traversal measure" is an offset measured from the initial node in the traversal to the traversal reference point. It is in the same units as the "weight" attribute of the links in the traversal.

Point Event. A zero-dimensional phenomenon that occurs along a traversal and is described in terms of its attributes in the extended database (see Figure 8).

Examples of point events include signs and accidents. Each point event in the linear referencing system data model has a "traversal measure" attribute. "Traversal measure" is an offset measured from the referenced traversal reference point to the point event. Point event traversal measures are in the same units as the traversal measures of the traversal reference points that they reference. A positive point event traversal measure expresses measurement in the direction of the traversal. A negative point event traversal measure expresses measurement against the direction of the traversal. Point events will typically have additional attributes in the extended database.

Linear Event. A one-dimensional phenomenon that occurs along a traversal and is described in terms of its attributes in the extended database (see Figure 8). Examples of linear events include pavement types, speed zones, and construction projects. Each linear event in the linear referencing system data model has "start traversal measure" and "end traversal measure" attributes that locate the linear event along the traversal. The traversal measures are offsets measured from the traversal reference points that they individually reference. Linear event traversal measures are in the same units as the traversal measures of the traversal reference points that they reference. Rules for direction of measurement are identical to those of point event traversal measures. Linear events will typically have additional attributes in the extended database.

Associations and Their Attributes

An **anchor section** goes *from* an **anchor point** *to* an **anchor point**. Each anchor section is associated with two, not necessarily distinct, anchor points in this way. Any number (0,N) of anchor sections can go from or to a given anchor point.

A **link** goes *from* a **node** *to* a **node**. Each link is associated with two, not necessarily distinct, nodes in this way. At least one link must go from or to a given node.

A **traversal reference point** must be *on* one and only one **traversal**. A traversal can have any number (0,N) of traversal reference points on it.

A **point event** must *reference* one and only one **traversal reference point**. A traversal reference point can be referenced by any number (0,N) of point events. A **linear event** must (1) *reference its start point* to one and only one traversal reference point and (2) *reference its end point* to one and only one traversal reference point. The traversal reference points that are referenced for the start and end of a linear event do not necessarily have to be distinct. A traversal reference point can be referenced by any number (0,N) of linear events.

A number of aggregation associations appear in the model. The **linear datum** *is composed of* at least two, not necessarily distinct, **anchor points** and at least one **anchor section**. A **cartographic representation** *is composed of* at least one **line**. A **network** *is composed of* at least one link and at least two, not necessarily distinct, nodes. A **linear referencing method** *is composed of* at least one **traversal** and at least one **traversal reference point**. A **traversal** *is composed of* at least one **link**. The links are ordered in this association, thereby, giving direction to the traversal (see Figure 6). A **link** *can be a component of* many traversals (see Figure 7).

A **cartographic representation** can *represent* zero or one **linear datum**. The linear datum can be represented by any number (0,N) of cartographic representations (see Figure 9). A **line** can *represent* any number (0,N.N) of **anchor sections**, including as many as two partial anchor sections (see Figure 10). An anchor section can be represented by any number (0,N.N) of lines, including as many as two partial lines. The lines are ordered and the association is assigned attributes to resolve the many-to-many and partial-object mappings. The association "*represents*," between line and anchor section, has "from position" and "to position" attributes. "From position" specifies the percentage of the first line in the list to be used as an offset from that line's start point to map the beginning of the anchor section onto that line. "To position" specifies the percentage of the last line in the list to be used as an offset from that line's start point to map the end of the anchor section onto that line.

A **network** is *referenced to* zero or one **linear datum**. The linear datum can have any number (0,N) of networks referenced to it (see Figure 11). A **node** can be *on* zero or one **anchor section**. An anchor section can have any number (0,N) of nodes on it (see Figure 12). A **link** lies *along* any number (0,N) of **anchor sections**. An anchor section can have any number (0,N) of links lying along it. The association "*along*," between link and anchor section, has a "directed sequence" attribute. The attribute specifies the sequence of anchor sections along which a link lies and the concurrence or contrariness of the direction of the link and the direction(s) of the anchor section(s). This association is required to eliminate ambiguities that can arise in its absence (Fletcher, 1995; Vonderohe and Hepworth, 1996).

A **linear referencing method** must *reference data to* at least one **network**. A network can have any number (0,N) of linear referencing methods that reference data to it (see Figure 13).

Clarification of Directionality

Figure 14 depicts the components of the model and a mapping that produces displayable coordinates for business data. Three object classes in the model have direction associated with them.

Each anchor section has direction, by definition. The direction of an anchor section could be initially established in any way, perhaps as a matter of convenience in the field. Anchor section direction is used to map lines and nodes onto anchor sections and to map anchor sections onto links.

Each link has direction because the order of the nodes it connects is specified. Link direction might be established according to the application the links will support. Link direction is used to map links onto anchor sections and to support network analysis.

Each traversal has direction, by definition. Traversal direction might by established by institutional factors (e.g., STH 10 South) or by analysis (e.g., pathfinding). Traversal direction is used to order traversal reference points and to map point events and linear events onto traversals.

Within the model, these three kinds of direction are reconciled by ordering and specifying "from/to" associations. Anchor section direction is established by specifying "from" and "to" anchor points. Link direction is established by specifying "from" and "to" nodes, then reconciled with anchor section direction by specifying a directed sequence of anchor sections when mapping links to anchor sections. Link direction is reconciled with traversal direction by ordering the links that compose the traversal.

Supported Operations

Using information from an implementation of this model, one can **locate** oneself in the field by first identifying which anchor section one is on (from a hardcopy cartographic representation or a listing of anchor sections, their anchor points, and the location descriptions of the anchor points). A measurement will then be made, along the linear facility, from the "from" anchor point, toward the "to" anchor point, to the unknown point, thus establishing its location.

A phenomenon (say, an accident) in the field can be **positioned** as a point event in the database merely by creating a record that identifies the traversal reference point to which the accident is referenced and specifies the traversal measure (+ or −) from the traversal reference point to the accident. The record will usually also contain values for other attributes of the point event (accident).

A point event in the database can be **placed** in the field by first using the traversal measure of the point event and the traversal measure of the associated traversal reference point to compute a cumulative offset from the initial node in the traversal to the point event. Then compute cumulative offsets from the initial node in the traversal to successive nodes in the traversal (using link weights) until a node is reached whose cumulative offset is greater than the cumulative offset of the point event. This node and the immediately previous node determine the link on which the point event lies. From the cumulative offsets, compute the offset of the point event from the "from" node of the link as a percentage of the weight of the link. Using the link/anchor section association and the node/anchor section associations determine the anchor sections and/or portions thereof that map to the link. Determine the length of the link from the distance attributes of the anchor sections. Compute the distance from the mapped location of the link's "from" node to the point event using the percentage offset and the length of the link in distance units. Determine the anchor section that contains the point event and compute the distance from that anchor section's "from" anchor point to the point event. Produce a hardcopy cartographic representation and print out the location description of the "from" anchor point. Discover the "from" anchor point in the field and, using the cartographic representation for direction reference, lay out the distance to the phenomenon represented by the point event.

A milepoint reference can be **transformed** into a project/engineering stationing reference on the same traversal by first comparing each project's "beginning of job" or "0+00" traversal measure to the traversal measure of the milepoint. If the milepoint is on any project, it will be on the project whose 0+00 has the largest traversal measure that is less than the traversal measure of the milepoint (here it is assumed that all project directions are the same as that of the traversal). For the selected project, determine if the "end-of-job" traversal measure is greater than the traversal measure of the milepoint. If so, the milepoint is on the project. Compute the offset from 0+00 to the milepoint. Express the result as engineering stationing.

ADDRESSING THE ISSUES

A brief synopsis of how the model addresses issues identified by workshop participants follows:

1. **Terms and Definitions.** A data dictionary is included with the model. Definitions from the literature, particularly the SDTS were used whenever possible. Terms having alternative meanings or interpretations in the transportation and GIS communities were avoided.

2. **Scoping.** The scope of the model is *linear* referencing. Primary spatial aspects of the model are topology in networks, distance measures in the linear datum, weights in networks, and offsets to locate zero-dimensional objects. The model includes cartographic representations in higher dimensions to provide coordinate references, to-scale visualization, and linkages to extended GIS data models. Issues such as polygonal representation of facilities at large scales and left/right offsets to off-facility features are not directly addressed by the model, but could be treated in extensions. Four fundamental operations are supported. Many other higher-level operations are also certainly supported. Temporal dimensionality remains unaddressed, except for a few ideas on versioning (see item 7, below). The development of robust space/time abstractions is an open research area.

3. **Schema Constructs.** The model was developed to be generic and to include the core

requirements of as many application areas as possible. The model is presented as an object model. Considering the sparseness of behavior included in the model, a relational form should be readily derivable. No hierarchies of classes are included in the model as presented. The model has components that are non-planar (networks) and components that are planar in many implementations (cartographic representations).

4. **Transformation.** Transformation is one of four fundamental operations demonstrated to be supported by the model. Transformations are possible between linear referencing methods, between cartographic representations, and between linear referencing methods and cartographic representations. Aggregation of events to support more generalized reporting should be supported by the model, although no demonstration is provided.

5. **Multi-Dimensionality.** The linear datum is the fundamental reference space for transformation, for linking to higher dimensions, and for solving the conflation problem. It is a topological object that includes distance measures. Proximity operations can take place at the linear datum level if distance is the desired metric, at the network level if link weight is the desired metric, or at the cartographic representation level if two- or three-dimensional analysis is desired. Use of GPS during data collection essentially provides a new cartographic representation of a linear facility for each pass in the field. Such data can be linked to the model by associating receiver positions with anchor points as they are encountered in the field.

6. **Methods and Coding.** The "lowest common denominator" is the linear datum. It consists of anchor points and anchor sections. Anchor points are zero-dimensional objects that must be unambiguously identifiable in the field. Therefore, a bridge cannot serve as an anchor point. The center point of a specified end of the deck of a specified bridge could serve as an anchor point. Starting and ending points of traversals are the appropriately ordered nodes of the first and last links in a traversal. The representation and referencing of ramps connecting roadways is not directly addressed by the model, but could be with extensions. Which geographic features are assigned external identifiers depends to some extent upon applications. The only features that require external identification are anchor points. Most applications would require identifiers for traversals and traversal reference points, at a minimum. The efficiency of the model for supporting operations on very large databases remains unknown. It should be remembered that the model is not intended as a specification. Refinements are possible.

7. **Data Integrity.** The model solves many data integrity problems arising from multiple copies of data. For example, linear referencing methods do not have to be explicitly imbedded in multiple cartographic representations. There is a single linear datum. Changes in the linear datum, caused by changes in alignment, generate a cascade of changes in the database (for mappings between the linear datum and cartographic representations, for mappings between the linear datum and networks, for specification of traversals, and for offsets of traversal reference points and events). Rules that associate these necessary changes must be developed. Temporal objects could be created through the use of version identifiers. With appropriate rules for assembly of temporal objects, they could be used to track changes over time.

8. **Institutional Policy**. The greatest incentive for policy concerning linear referencing systems is cost savings realized from data integration, data sharing, and reduction of chaos. Some agencies have already adopted policy in this regard (e.g., Utah). Use of a linear referencing system must be simple and straightforward. Even so, not all users must understand the use of all methods. Easily understood procedures must be developed for use in both the field and the office.

ASPECTS OF IMPLEMENTATION

Anchor Points and Intersections

Anchor points are not necessary at even-valence intersections. Such intersections can contain anchor points, but they are not required to. Any two anchor sections can cross without sharing an anchor point. Each odd-valence intersection must include an anchor point. However, all but one anchor section can pass through such an intersection without including the anchor point. This characteristic reduces the number of linear datum objects to be developed, managed, and maintained. It also places a lower bound on the number of anchor points required in the linear datum. The minimum number is the sum of the number of termini and the number of odd-valence intersections. However, other design considerations, such as the required accuracy of the underlying system of measurements, can cause the number of anchor points to increase beyond this theoretical minimum.

Bi-Directional and Multi-Lane Facilities

Choices must be made for representing bi-directional facilities. At the network level, individual links represent each direction of travel. The issue is the number of required anchor sections. A single roadway, having two directions of travel, can be represented by a single anchor section. This choice results in a minimum number of datum objects and has the further advantage that any individual object in the field (real world), serving as a reference point for traversals in both directions, has a single datum location. An example of such an object is the intersection of a roadway centerline with an edge of a bridge abutment. With this choice, spatial distinctions among events must be maintained at the link or traversal level. Otherwise, events of the same kind, which should be associated with opposite directions of travel, risk having the same offset along the single anchor section.

A single roadway, having two directions of travel, can be represented by two anchor sections (one for each direction). This choice has the advantage that spatial distinctions among events can be maintained at the linear datum level. Disadvantages include doubling the number of anchor sections and developing and maintaining two linear datum addresses for any single object in the field that serves as a reference point for traversals in both directions.

Two anchor sections are necessary in the unlikely occurrence of a single roadway with opposite traveled ways whose lengths differ by an amount large enough to require their resolution. A similar statement can be made for representation of multiple lanes in a single direction. Information on events at the lane level is useful for construction and emergency vehicle routing, signal timing, congestion management, and other operational functions and ITS applications. Fohl et al. (1996) propose representing roadway centerlines as continuous linear spatial objects with start and stop points for lanes located by offsets along them. They argue that theoretical and technological limitations on the accuracy of absolute positioning obviate the need for explicit spatial representation of lanes. In fact, the potential accuracy of relative measures, such as lane length, might be great enough to make metric distinctions among lanes, but the need for doing so from an application perspective probably does not exist.

A divided highway with separate roadways for each direction of travel requires an anchor section for each roadway. The anchor sections terminate at distinct anchor points and often have different values for the distance attribute. Distinct objects in the field serve as reference points for traversals along the individual roadways.

Scalability

Scalability is traditionally an issue of cartographic representation and how spatial abstractions are affected by display at different scales. Much research has been done on map generalization and methods for transforming maps from larger to smaller scales. A linear referencing system has a certain level of abstraction, but it has no single cartographic scale. In fact, it has as many cartographic scales as there are cartographic representations with distinct source scales linked to it. This aspect of scalability of a linear referencing system is manifested in the one-to-many association between linear datum and cartographic representation.

A divided highway, represented by two anchor sections, might appear as a single line on a small scale map and as two lines on a large scale map. The anchor sections would both be associated with the single line for display at small scale. Each of them would be associated with an individual line for display at large scale. Two point events, each with a linear datum address in a separate anchor section, would appear to be co-linear at small scale and in different lines at large scale. The Wisconsin DOT has an effective demonstration of this characteristic of the Ries (1993) link/site model (Ries, 1995).

Zero-dimensional objects in the linear referencing system data model (e.g., anchor points and traversal reference points) are zero-dimensional objects on the ground (e.g., intersections of centerlines). Distances in the data model are distances as measured on the ground. Spatial abstraction takes place at the one-dimensional level. Roadways, pavements, and many of the things that happen along them (events) are two- and three-dimensional, yet they are represented as having linear locations. The second and third dimensions are sometimes represented as attributes (e.g., pavement width and thickness). In this manner, the linear referencing system is similar to a surface model that represents elevation as an attribute of horizontal location.

Should single roadways with travel in both directions be modeled by one anchor section and two anchor points, two anchor sections and two anchor points, or two anchor sections and four anchor points? The answer lies in what we are truly trying to represent and what level of the model we select for maintaining spatial distinctions. The former must be addressed from an application perspective. The latter must be addressed from a data management perspective.

SUMMARY

A generic data model for linear referencing systems has been developed. The model includes multiple cartographic representations, multiple networks, and multiple linear referencing methods to which any amount of business data can be tied. The model supports integration of attributes attached to various spatial databases without requiring registration of cartographic representations in coordinate space. Instead, they are all linked to a single common linear datum.

The model supports a set of fundamental operations that links the database world and the real world and allows transformations within the database world. The model will also support network analysis and basic GIS operations, although examples have not been developed.

The model is intended as a description of the core common requirements of as many application areas as possible. A need for extension of the data model to include particulars for specific application areas should be expected. Potential application areas include infrastructure management, transit, freight, ITS, urban planning, waterway navigation, and seismological testing.

Developments Since Publication of Initial Results

Since the workshop, interest and activity associated with linear referencing systems has increased. FHWA is supporting development of a "best practices" manual for linear referencing. The ITS community is striving for standards and a generic data model for linear referencing (Siegel et al., 1996). The GIS-T Pooled Fund Study Team adapted the data model reported herein during Phase B of their research (Fletcher, 1995) and developed a linear referencing engine as proof-of-concept for the data model. A number of standards efforts include consideration of linear referencing (Hickman, 1995; Scarponcini, 1995). Systems have been developed to address the conflation problem (Siegel, 1995; Brown et al., 1995).

Implementation issues have been examined (Sutton and Bespalko, 1995). The workshop data model has been extended (Dueker and Butler, 1997). A methodology for design of the field component of linear referencing systems has been developed (Vonderohe and Hepworth, 1996). A call has come for development of a unified linear referencing system with a common linear datum to support the transportation and navigational data needs of civilian government, the military, and the private sector (Fletcher et al., 1996). BTS has hosted a World Wide Web-based discussion of linear referencing issues for the GIS-T Task Force of the Transportation Research Board. A dialog between Wende O'Neill of Utah State University, Bruce Spear of BTS, and the principal author of this digest serves as the concluding section of the digest.

DISCUSSION AND CLOSURE

The following questions and comments from O'Neill and Spear and answers and comments from Vonderohe are intended to clarify and elaborate on aspects of the linear referencing systems data model, to address issues raised subsequent to publication of the workshop report, and to explore some practical aspects of implementation of the model.

Question 1: According to the definition of "anchor point," forms of location descriptions can vary and be quantitative or descriptive or both. If we use only a named road description for an intersection, then it is likely that some anchor points have the same description. Is this a problem? Don't we have to attach additional information, such as a coordinate point or a direction (N,S,E,W)?

Answer: Yes, it is certainly a problem if we use only named roads in most cases. According to its definition, an anchor point must be uniquely identifiable and recoverable in the field. A location description that does not fulfill this characteristic is insufficient and cannot be used. No description is really adequate for actual implementation unless it clearly and unambiguously defines the location of a point for people using the description in the field. Anchor points are actually points, so we should probably be saying something like "The intersection of the centerlines of the traveled ways at the southwestern most of two intersections of Bonnie Branch Rd. & Bonnie Acres Dr." We make measurements of distances between these points, so they must be points.

Question 2: How do we define anchor sections? Using the named route notion (all road segments with the same name belong to a route) and the concepts described above to define anchor sections, design of the anchor sections ensures that an anchor point is never the "from" anchor point for more than one anchor section. The definition of node implies knowledge of the anchor section for which the linear datum measurement is relevant. However, it should be explicitly stated that if anchor points are not uniquely used as "from" points on a single anchor section, then the node records must contain explicit reference to the anchor section. This solves the multiple path problem.

Answer: There is no need to keep anchor points from being the "from" anchor point for more than one anchor section. Anchor points and anchor sections are not restricted in this way. Two or more anchor sections can share a common anchor point—and that anchor point can be the "from" anchor point for all of them (but it doesn't have to be; it could be the "to" anchor point for some of them). Anchor sections can have common anchor points, but they cannot overlap. The direction of each anchor section is totally arbitrary. That is because their sole purpose is to serve as the basis for locating other things.

The model requires explicit reference of nodes to anchor sections because there would be ambiguity without it. The object model diagram in Figure 2 includes a direct association between nodes and anchor sections. This would manifest itself as a pointer from a node to an anchor section. Since the anchor section knows which anchor point is its "from" anchor point, the node can be located along the anchor section.

Question 3: In designing the linear datum, you first define your points, then the lines connecting them correspond to existing road facilities. With a network, you first define your connections (flows of interest), then the points are self-evident.

Answer: The linear datum is similar to a network in some ways, but it is different in at least one very important way. Flows are of no concern in designing the datum. The datum does not tell us how to get from "here" to "there" like a network does. It just tells us where things are. There are no routes or traversals in the datum.

Each anchor section represents a single real object (e.g., a section of roadway) and that real object is represented by a single anchor section. Any real thing has one and only one true location at any point in time (unless it is a quantum particle or something). The purpose of the linear datum is to provide a basis for describing the unique locations of things.

Question 4: Using anchor sections with calibrated distances between each adjacent pair of intersections results in higher accuracy in placing linearly referenced events along a cartographic line. Basically, the error is spread over a shorter distance. Anchor sections that stretch from one county boundary to another or one state boundary to another (like I-15 between the Idaho border and the Arizona border), with no calibration points in between, will result in substantial error in trying to place a linear event. Accuracy in placing events along a centerline will be more of a problem in rural areas, where the distance between anchor points is greater, and on very curvy roads. However, the extreme opposite representation, a calibrated anchor section between each pair of intersections, brings back the multiple route problem whose solution is to reference the anchor section with a name descriptor.

Answer: Two important points of discussion are raised above:

1. Do we have to agree on names of roads in order to identify anchor sections? I vote a resounding "No"! What we need to agree on is a set of neutral identifiers—one for each anchor section. Then, everyone can associate whatever names they want

with the identifiers (and, thereby, the anchor sections). This way, if the Municipal Manager wants to call the roadway in front of my house "Main Street" and the DOT wants to call it "Highway 51," they can both have their way and still be able to share data. The events that they are interested in are happening along the roadway—and there is only one roadway, even though they refer to it by different names. Those events of interest will be located by offsets along the same individual anchor sections representing that roadway—for both the Municipal Manager and the DOT. Road names and routes (traversals) change too often and different parties and authorities use different ones of them to identify the same thing.

2. What is the optimum length of an anchor section? This is a question that I think should be vigorously debated in order to identify all the factors and the weights that should be given to each of them as they contribute to the answer. It is one of a few questions at the heart of the linear referencing system design problem. Another, for example, is "What is the required accuracy of anchor section distance measures?" The answers to these, and a few other key questions determine the optimum configuration of the linear referencing system in the field. By "optimum," I mean the least-cost configuration that meets the accuracy requirements of its users. A method for finding the answers to these questions is what the report to Sandia Labs is about (Vonderohe and Hepworth, 1996).

I will now put in my two cents' worth: I think the factors that are most important in addressing this question have to do with linear referencing and not cartographic representations. When we do linear spatial analysis (e.g., routing, allocation, trip assignment), we are concerned with the accuracies of linear locations, not x,y locations. (As an aside, I believe that the primary reason we need a linear referencing system is to do linear spatial analysis—I mean this in a broad sense—not just pathfinding, etc.). For example, a pavement management application can be very effective by operating only on linear data, without ever referencing two- or three-dimensional data. So the accuracy needed here is in linear locations. Now, if the pavement manager wants to display the results on a map, we need to be concerned with the accuracy of relationships between linear and cartographic data for visualization, not analytical purposes. This does not mean that accurate visualization is not important. Also, there are other kinds of applications that require combinations of linear and cartographic or x,y,z data for analytical purposes (e.g., linking of linear and area data, such as wetlands; linking of linear and x,y,z data for vehicle navigation). In any case, it is possible to enforce constraints on linear referencing system design so that linear locations will have accuracies that are compatible with any selected map scale. Anchor section lengths and the accuracies of on-the-ground distance measurement technologies combine to set an upper bound on the accuracies of linear locations. The design solution turns this relationship around and transforms the accuracies of linear locations (specified according to user requirements) into specifications for anchor section distances and measurement procedures.

Question 5: Why does a traversal have to be made up of whole links? Why can't a traversal cover part of a link?

Answer: The notion of traversals being made up of whole links arose from the Milwaukee workshop. Excerpts, transcribed from tape recordings of the workshop, include the following: "traversal is ordered set of whole links" and "traversals begin and end at nodes." The definition of traversal was drawn from these notes and others that were taken at the workshop. In my cover memo to workshop participants, concerning the first-draft report, I asked for feedback on this issue. The memo placed the issue under "points of contention and open questions." The question was put as "Should traversals be allowed to start/stop in mid-link?" I received explicit feedback on this question from three participants. The first said "No" and gave a number of reasons why. The second said "Yes, but not a strong preference...." The third said "We need to define a 'start' and a 'stop' on a link." Based upon this response, I made no change. All the reviewers that responded to the first-draft report were given the opportunity to respond to a few questions I had before preparation of the second-draft report.

I personally do not feel strongly about the modeling question at the heart of this issue. Traversals go from someplace to someplace and "someplace" is usually represented in a network as a node. But if we let traversals start/stop in mid-link, the model is more flexible. I think that this central issue should perhaps be debated further, but the model, as it now stands, best represents the sentiments of the Milwaukee workshop participants as well as I could determine those sentiments.

Question 6: Would you please give a real-world example of the link/anchor section ambiguity problem and the need for the association between links and anchor sections in the data model. When would an anchor section form loops (particularly in light of the previous discussion on anchor section design)?

Answer: My original thinking on the link/ anchor section ambiguity problem was that it would not arise very often in real situations. Since then, I have come to realize that it is much more prevalent. In fact, it occurs almost all the time. There is ambiguity any time there is more than one sequence of anchor sections between the two nodes. This same ambiguity, and the same means for resolution, were discovered earlier by David Fletcher and others on the GIS-T Pooled Fund Study when the NCHRP model was being prototyped in the Linear Referencing Engine.

NOTE: Here "sequence of anchor sections" sounds like "path," but I am using "sequence of anchor sections" to keep from implying that the linear datum is a network. Anchor sections can share anchor points, but the linear datum does not have to be connected in the way that a network is.

If I say merely that for link 10, its "from" node is on anchor section 5 at offset 0.54 and its "to" node is on anchor section 5 at offset 1.32, this does not mean that link 10 is entirely on anchor section 5. The link could lie on any sequence of anchor sections from the "from" node to the "to" node. Each of the possible alternative sequences begins and ends on anchor section 5. There are two solutions to this problem. One of them is enforcement of a rule that restricts links to lie on one and only one anchor section. I think this is too restrictive. There are applications, such as freight, that might use networks with links that span anchor points. The other solution to the problem is to require a link-to-anchor section association, so that the anchor section(s) on which the link lies are explicitly stated. This leaves the model flexible and supporting as many applications as possible, which was one of the original goals of the Milwaukee workshop participants. This is the solution originally conceived by the GIS-T Pooled Fund Study.

It is also possible for a single anchor section to form a loop. There is no rule preventing the "from" anchor point and the "to" anchor point of a given anchor section from being the same. Circle drives can be modeled in this way. However, the direction (clockwise or counter-clockwise) around the loop must be specified. This is analogous to the cul-de-sac addressing problem identified by Sutton and Bespalko (1995). The solution is the same. Either we model it directly, or we agree on a standard.

Question 7: The pathologies addressed by Sutton and Bespalko (1995) are related to implementation. They indicate that discontinuous traversals are problematic and require a set of rules for coding. For example, should they be referenced as continuous or should traversal measures be re-initiated at each break?

Answer: There is a decision to be made at implementation time, as you point out in your comment below. Traversals do not have to be continuous, but if they are not, we must decide whether or not to reference them as if they were continuous.

Question 7 continuation: So if you have a traversal reference point A that is at offset oA on anchor section 1 from anchor point 1 and traversal reference point B that is at offset oB on anchor section 5 from anchor point 32, how do you know if there are continuous traversal segments between these two traversal reference points or not? You would have to identify the links that make up the traversals, and using the nodes for the links determine if there is a connected path between the traversal reference points or not.

Answer: Yes, that is exactly how I would find out if a traversal was continuous between two traversal reference points.

Question 7 continuation: The other part of this question is what measure do I give to traversal reference point B (which is not oB since this is the offset along the anchor section and not along the traversal) when the traversal is discontinuous. Do I reference it as zero to indicate the beginning of a new section in the traversal or as a positive offset from the "true" beginning of the traversal even though there will be measured pieces that are not part of the traversal? This is an implementation question and can only be addressed in the context of the software and applications. (I cannot really think of a situation where discontinuous route measurement is an issue. For example, if I build snow plow routes, I can indicate blade-up/blade-down sections along that route and calculate total plowed miles by summing the lengths of the blade-down condition. Or, for transit, we can have in-service/out-of-service conditions along the traversal and calculate total miles in service. In other words, we reference events to traversals or to the linear datum and work through the logical transformations desired.)

Answer: I wholeheartedly agree that this is not really an issue. We just have to decide how to reference noncontinuous traversals. In one possible database implementation, all locations of interest (e.g., events, traversal reference points, nodes) would be stored as offsets along anchor sections. Queries would invoke transformation routines that would generate appropriate traversal offsets, in whatever linear referencing method is called for, on the fly. Inverse transformations would determine anchor section offsets for data, collected by whatever linear referencing method, coming in from the field.

Question 8: The overall data model is designed to facilitate data integration over multiple cartographic databases and for multiple network representations with presumably many linear referencing methods. Do you have an estimate of how many DOTs actually maintain multiple cartographic databases and multiple network representations?

Answer: 5 or 6 years ago, during the initial survey stage of the NCHRP work, a number of agencies were using different maps at different scales to build their GIS spatial databases. Of course, this violates some basic rules of cartography. They were doing this for at least two reasons: (1) this mix of maps was all that was available and (2) the products from vendors were based upon data models that allowed only one cartographic representation. Therefore, the scales had to be mixed.

I do not have comprehensive current information on what the agencies are doing with multiple maps and multiple networks today. I suspect that many have digital maps (if not spatial databases) at different scales (even though they might not be trying to actively maintain all of them). I also suspect that they have tremendous difficulty relating their linearly referenced data to more than one network and more than one cartographic representation without having multiple copies of their linearly referenced data. Software in place today just does not support this kind of multiplicity.

Perhaps we should view the NCHRP model as a way of building many integratable databases instead of one great big integrated database. In this way, it supports data sharing, not only within a single organization, but also across organizational boundaries. If we all had the same linear datum, the sharing could be not only horizontal but also vertical. The metropolitan planning organization (MPO) could give its linearly referenced data to the DOT and vice versa. The DOT could analyze the MPO's data, integrate it with the DOT's data, and display the results, using the DOT's linear referencing methods, networks, and cartographic representations. The MPO could analyze the DOT's data, integrate it with the MPO's data, and display the results, using the MPO's linear referencing methods, networks, and cartographic representations. This could work in the same way that the DOT currently gets the MPO's two-dimensional cartographic representation of streets in state plane coordinates and transforms it into Universal Transverse Mercator (UTM) coordinates to match those of the DOT's cartographic representation. They can do this because there is a common geodetic datum for two-dimensional location referencing.

Question 9: How is the linear datum similar to a network? How is it different? Can the linear datum be used as a network?

Answer: In some ways, the linear datum is similar to a network. However, there are important differences. The linear datum is not fully connected in the way a network is. For example, anchor sections can pass through intersections without encountering an anchor point. It might be possible, but it is probably highly unusual, for a link to pass through an intersection without encountering a node (the node, of course, breaks the link at the intersection). The reason that anchor sections can behave this way is that they have nothing to do with routing, pathfinding, turns, stops, impedances, or how to get from here to there. Their sole purpose is to represent transportation facilities in such a way that unique locations can be established for things of interest. Nodes represent intersections and termini for navigation through a network. Anchor points do not. Anchor points represent appropriate places for endpoints of anchor sections for the purpose of location referencing. Many times, an anchor point and a node will have the same location. But other times they will not. The only places where they are *required* to have the same location are at odd-valence intersections and termini. As an example, it might be necessary to place an anchor point (but not a node) along a long stretch of roadway to create two anchor sections where there is only one link. This can happen because of accuracy considerations in the design of the linear referencing system.

The only reason that anchor sections share anchor points is to reduce the number of anchor points that have to be managed. There is no functional reason beyond simplification of the linear datum—but that is a pretty good reason. It is not necessary to give multiple anchor point identifiers to a single real-world feature just because it is an anchor point for more than one anchor section. I suppose it is possible to generate a number of link/node subnetworks from the anchor points and anchor sections of the linear datum, but I do not think such subnetworks would be very useful for the kinds of things we need to do with networks. The optimum configuration of the linear datum, designed for location referencing, and the optimum configuration of a corresponding network, designed for wayfinding, are not the same. I think that, even though the datum is similar to a network in some ways, it should not be thought of as a network. They are really different beasts.

Question 10: Anchor points have a description associated with them to help locate them in the field. Anchor sections have no description other than the two anchor points at their ends, yet their real-world lengths are also recoverable in the field. How can you determine the length of an anchor section if you are located at one anchor point and there are multiple paths to the other anchor point? It seems to me you need some identifier for the anchor section as well.

Answer: We certainly need identifiers for anchor sections. If these are neutral identifiers (as I think they should be), they might not, unto themselves, give us enough information to locate the anchor section in the field. However, the identifiers will be associated with whatever names various authorities assign to roadways (perhaps multiple names). With the associated names, we can find them in the field. If we associate the anchor sections with cartographic representations, we can also make a map and carry that into the field.

Here is a related issue. Are the names (not the identifiers) we assign to roadways different from the names we assign to traversals? I think they are very similar, if not the same. "Highway 51" and "Main Street" are names for roadways. Are not these close to the names we would assign to traversals? Maybe for traversals we would add something on direction, like "Highway 51 North."

Question 11: Why is it that cartographic representations can consist of disjoint groups of lines?

Answer: My recollection of the intent of the Milwaukee workshop participants was to enable partial two- or three-dimensional mapping. That is, we can still have a linear referencing system even though we might not have a map for all of it (or even part of it for that matter—we can have a linear referencing system without having any map of it at all).

Question 12: Why do you exclude chains from the network component of your data model? Is it because you want to use links and nodes to describe a purely topological object, while chains mix topology and geometry?

Answer: Yes, chains are excluded from networks for that purpose. Chains are included as a possible kind of line making up a cartographic representation. This addresses a problem that some folks have been struggling with for a few years. These people are the ones trying to link transportation planning models to GIS databases and software. The networks in the planning models and the spatial databases in the GISs have nodes in different places. Network nodes exist for analytical purposes. GIS spatial database nodes (at the ends of chains) often exist because of the way in which the data were captured off a hardcopy map. If we had a linear datum, we could link both the planning network and the GIS spatial database (cartographic representation) to it. Then it would not matter if the nodes used in the network and the nodes used in the GIS spatial database were the same nodes. Network analysis would be done in the network and the results would be displayed in the GIS spatial database (cartographic representation). It is very true that we would still need to go to the effort of linking the network and the cartographic representation to the linear datum. But everyone would have the same linear datum. So, the next time we wanted to include some other (or somebody else's) cartographic representation or network, it would already be linked to the linear datum, or, at worst, if it was not already linked to the linear datum, we would only have to link that object to the linear datum instead of to every cartographic representation and network we already had.

Question 13: Based on your definitions, traversals, point events, and linear events, are all built upon links. Does this mean that you cannot locate any linearly referenced events without first building a link/node network?

Answer: No, it does not mean that. As discussed above, if I were implementing a database, I would store locations of things of interest as offsets along anchor sections. Therefore, we can know their locations without having a network. However, if we want to go further and do spatial analysis, even simple things, like finding the distance between two events on anchor sections that cannot be connected through the linear datum (because the linear datum does not have to be connected), then we need a network that is appropriately connected. Also, the locations of new events, coming in from the field, will be referenced by some method (e.g., county-route-milepoint). We need a network to transform that method-based reference into a linear datum-based reference. But if we have a method for referencing, then we automatically have a network.

BIBLIOGRAPHY

Brace, K. and B. Peterson, "Cook County, Illinois TIGER/Line Conflation and Update Project," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Albuquerque, NM (1994).

Briggs, D.W. and B.V. Chatfield, "Integrated Highway Information Systems," *NCHRP Report 133*. TRB, Washington, DC (1987).

Brown, J.N., A.L. Rao, and J. Baran, "Automated GIS Conflation: Coverage Update Problems and Solutions," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Sparks, NV (1995).

Cihon, R., *Geographic Information Systems Resource Guide*, Washington State Department of Transportation, Olympia, WA (1996).

Clark, J. and L. Bain, "Integrating a Transportation Model Network into Enhanced TIGER/Line," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Albuquerque, NM (1994).

Deighton, R.A. and D.G. Blake, "Improvements to Utah's Location Referencing System to Allow Data Integration," presented at the Transportation Research Board 3rd International Conference on Pavement Management, San Antonio, TX (1993).

Dueker, K.J., "Geographic Information Systems and Computer-Aided Mapping," *Journal of the American Planning Association*, Vol. 53, No. 3 (1987).

Dueker, K.J. and J.A. Butler, "GIS-T Enterprise Data Model with Suggested Implementation Choices," Center for Urban Studies, School of Urban and Public Affairs, Portland State University, Portland, OR (1997).

Dueker, K.J. and R. Vrana, "Dynamic Segmentation Revisited: A Milepoint Linear Data Model," *Journal of the Urban and Regional Information Systems Association*, Vol. 4, No. 2 (Fall 1992).

Federal Highway Administration, *Highway Performance Monitoring System Field Manual*, Office of Highway Information Management, FHWA Order M5600.1B (1993).

Fletcher, D.R., "Modelling GIS Transportation Networks," *Proceedings Vol. II*, Urban and Regional Information Systems Association, 25th Annual Conference, Fort Lauderdale, FL (1987).

Fletcher, D.R., *GIS-T Pooled Fund Study Phase B Summary Report*, Alliance for Transportation Research, Albuquerque, NM (1995).

Fletcher, D.R., J. Espinoza, R.D. Mackoy, S. Gordon, B. Spear, and A.P. Vonderohe, "The Case for a Unified Linear Referencing System," *Proceedings*, Workshop on Enterprise Location Referencing Systems: Policies, Procedures and Standards for Implementation, Salt Lake City, UT (1996).

Fohl, P., K.M. Curtin, M.F. Goodchild, and R.L. Church, "A Non-Planar, Lane-Based Navigable Data Model for ITS," National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA (1996).

Hickman, C., "Feature-Based Data Models and Linear Referencing Systems: Aids to Avoid Excessive Segmentation of Network Links," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Sparks, NV (1995).

Mapping Science Committee, *Toward a Coordinated Spatial Data Infrastructure for the Nation*, National Academy Press, Washington, DC (1993)

Michigan Department of Transportation, *1984 Reference System Report*, Database Task Force, Reference System Review, Lansing, MI (1984).

National Cooperative Highway Research Program, "Highway Location Reference Methods," *Synthesis of Highway Practice 21*, TRB, National Research Council, Washington, DC (1974).

Nyerges, T.L., "Locational Referencing and Highway Segmentation in a Geographic Information System," *ITE Journal* (March 1990).

Nyerges, T.L. and K.J. Dueker, *Geographic Information Systems in Transportation*, Office of Planning HPN-22, Federal Highway Administration, U.S. Department of Transportation, Washington, DC (1988).

Peterman, V., "Conflation: Merging Spatial Data Sets," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Albuquerque, NM (1994).

Ries, T., "Design Requirements for Location as a Foundation for Transportation Information Systems," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Portland, OR (1993).

Ries, T., "Integrating Governments for Transportation Purposes Using a Geospatial Framework," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Sparks, NV (1995).

Rowell, R., "Theory and Practice: Linear Referencing at the Idaho Transportation Department," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation," Kansas City, MO (1996).

Rumbaugh, J., M. Blaha, W. Premerlani, F. Eddy, and W. Lorensen, *Object-Oriented Modeling and Design*, Prentice Hall, Englewood Cliffs, NJ (1991).

Scarponcini, P., *Location Data Modeling Effort Final Report*, prepared for Minnesota Department of Transportation, Graphic Data Systems Corporation (1994).

Scarponcini, P., "A Method for Determining a Standard Linear Referencing Scheme," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Sparks, NV (1995).

Siegel, D., "A Non-Traditional Solution to Conflation," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Sparks, NV (1995).

Siegel, D., C. Goodwin, and S. Gordon, *ITS Datum Prototype Final Design Report Task C: Spatial Data Interoperability Protocol for ITS Project*, Federal Highway Administration, IVHS Research Division, McLean, VA (1996).

Sutton, J., and S. Bespalko, *Network Pathologies: Phase 1 Report*, Sandia National Laboratories Transportation Systems Analysis GIS Project, Document No. AH-2266, prepared by GIS/Trans, Ltd. (1995).

United States Geological Survey (USGS), *Spatial Data Transfer Standard: Part 1 Logical Specifications*, Reston, VA (1992).

Vonderohe, A.P., L. Travis, R.L. Smith, and V. Tsai, "Adaptation of Geographic Information Systems for

Transportation," *NCHRP Report 359*, TRB, National Research Council, Washington, DC (1993).

Vonderohe, A.P., C.L. Chou, F. Sun, and T. Adams, "Results of a Workshop on a Generic Data Model for Linear Referencing Systems," *Proceedings*, AASHTO Symposium on Geographic Information Systems in Transportation, Sparks, NV (1995).

Vonderohe, A.P. and T.D. Hepworth, "A Methodology for Design of a Linear Referencing System for Surface Transportation, Final Report," Project AT-4567, Sandia National Laboratories (1996).

Wisconsin Department of Transportation, *Location Control Management Manual*, Madison, WI (1996).

ACKNOWLEDGMENTS

The commitment and effort of the workshop participants is gratefully acknowledged. This report is a result of their work. The participants and their affiliations at the time of the workshop were:

Participant, Affiliation

Teresa Adams, University of Wisconsin—Madison
Hillary Armstrong, Sandia National Laboratory
Mark Bradford, Federal Highway Administration
Jim Carroll, Michigan DOT
Richard Church, Univ. of California—Santa Barbara
Chih-Lin Chou, University of Wisconsin—Madison
Ron Cihon, Washington DOT
Don Diget, Michigan DOT
Charles Dingman, United States Bureau of the Census
Ken Dueker, Portland State University
Joe Ferreira, Massachusetts Institute of Technology
David Fletcher, GeoDigm Consultants, Inc.
Cecil Goodwin, University of Tennessee
Stephen Gordon, Oak Ridge National Laboratory
Edward Granzow, Urban Analysis Group
Robert Harris, GIS Trans
Charlie Hickman, United States Geological Survey
Dale Honeycutt, ESRI
Wen-Jing Huang, University of Wisconsin—Madison
Peggi Knight, Iowa DOT
Roy Larson, Minneapolis/St. Paul Metro Council
Simon Lewis, GIS Trans
Frank Lockfeld, Center for Urban Analysis
David Loukes, GeoPlan
Bill McCall, Iowa State University
Daniel McHugh, MTA New York City Transit
Alisoun Moore, Maryland State Highway Administration
Tim Nyerges (Facilitator) University of Washington
Paula Okunieff, Viggen Corporation
Ken Opiela, Transportation Research Board—NCHRP
Tom Ries, Wisconsin DOT
Jay Sandhu, ESRI
Bill Schuman, Intergraph
Susan Scott, SEI Technology
Robert Smith, University of Wisconsin—Madison
Bruce Spear, U.S. Department of Transportation
Todd Stellhorn, ESRI
Jung-Gon Sung, University of Wisconsin—Madison
James Tucker, GDS Corporation
Alan Vonderohe (PI), University of Wisconsin—Madison
Kirk Weaver, New Jersey DOT
Lynn Williams, Intergraph

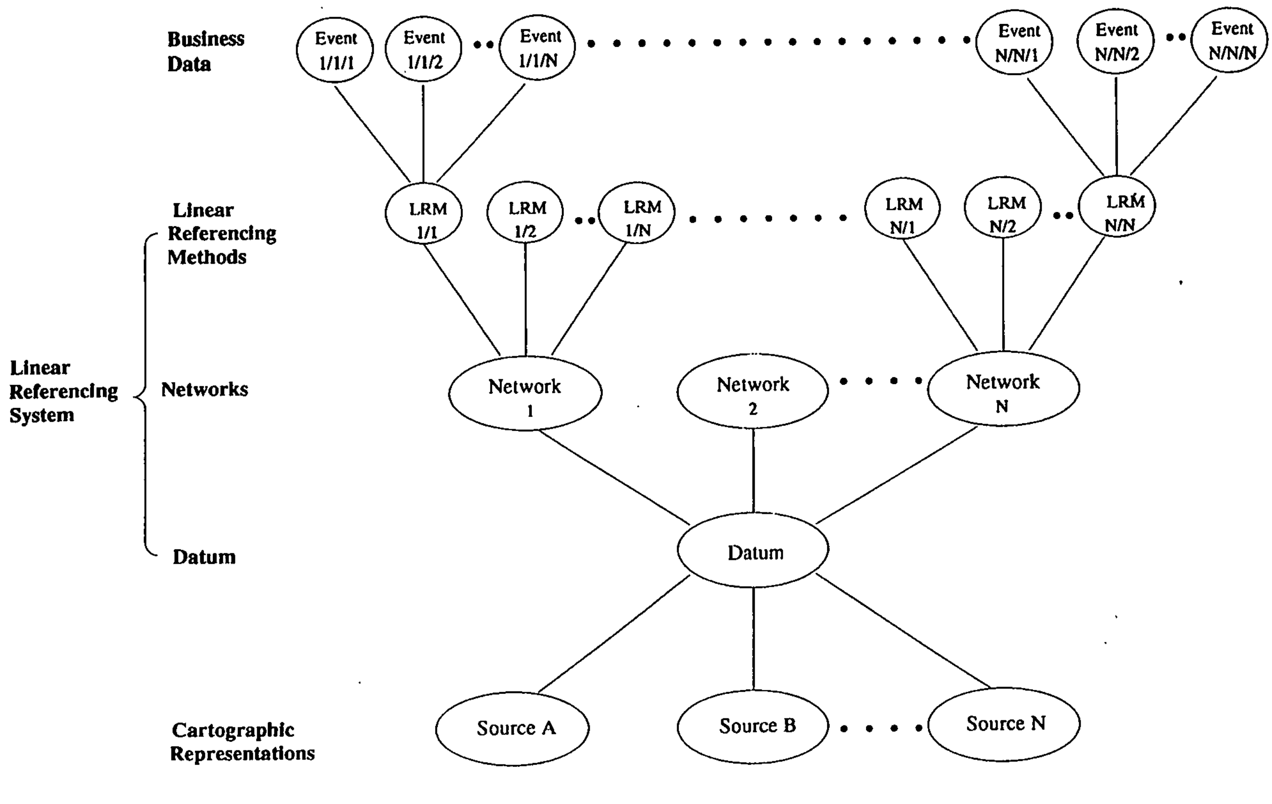

Figure 1. Model overview.

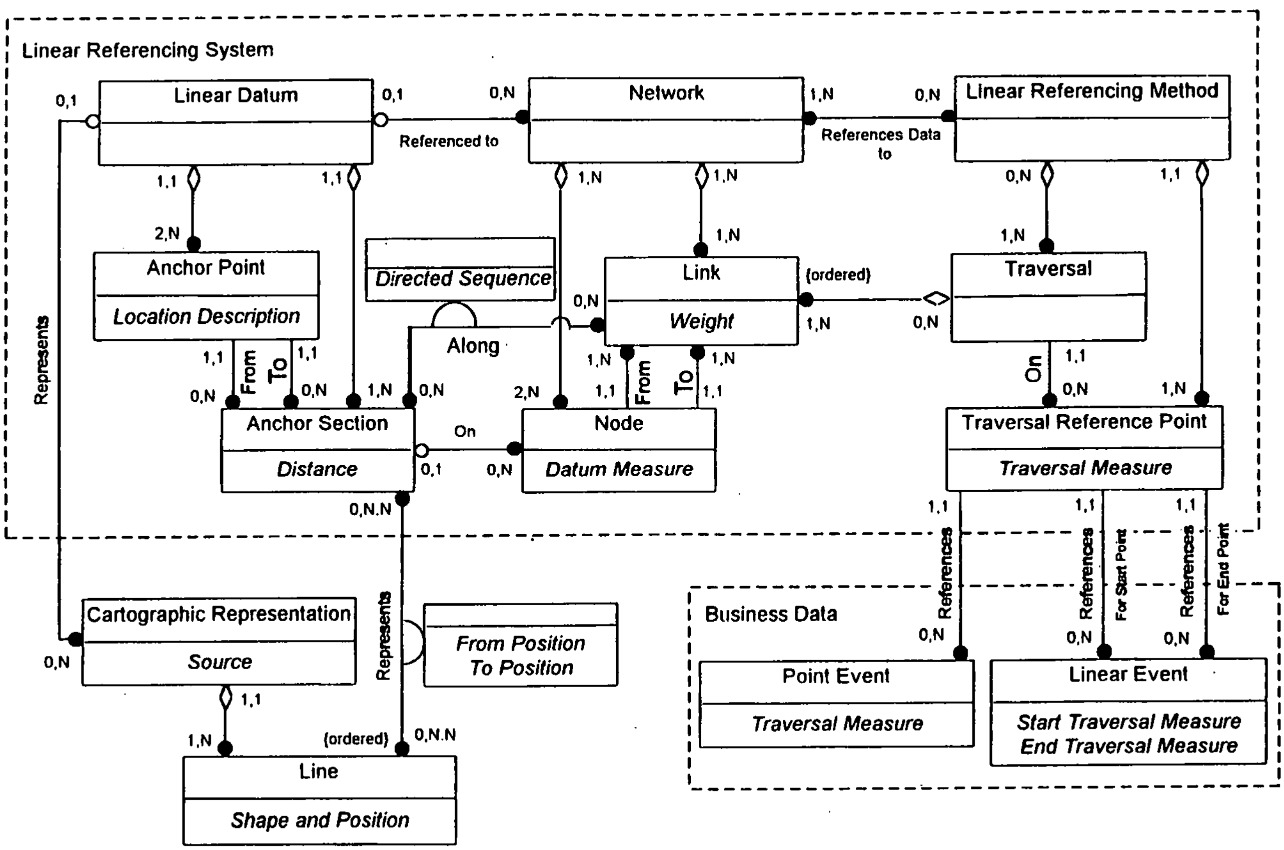

Figure 2. The linear referencing system data model.

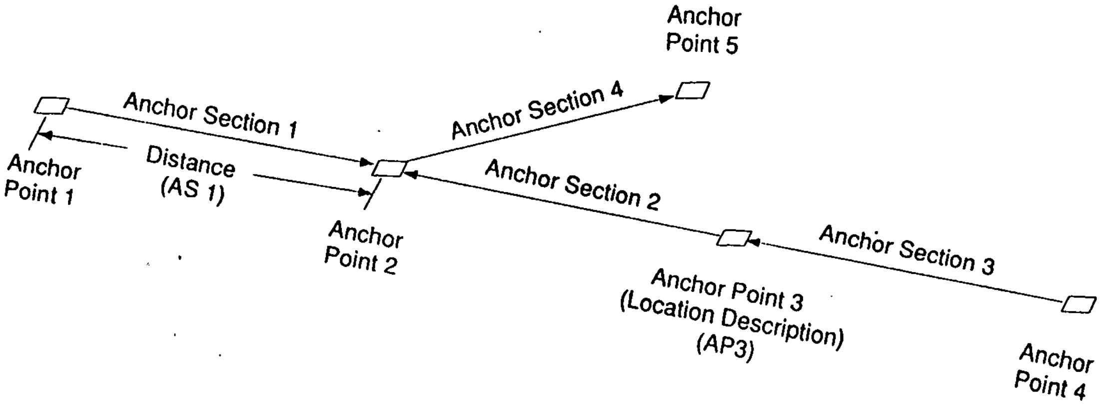

Figure 3. Datum.

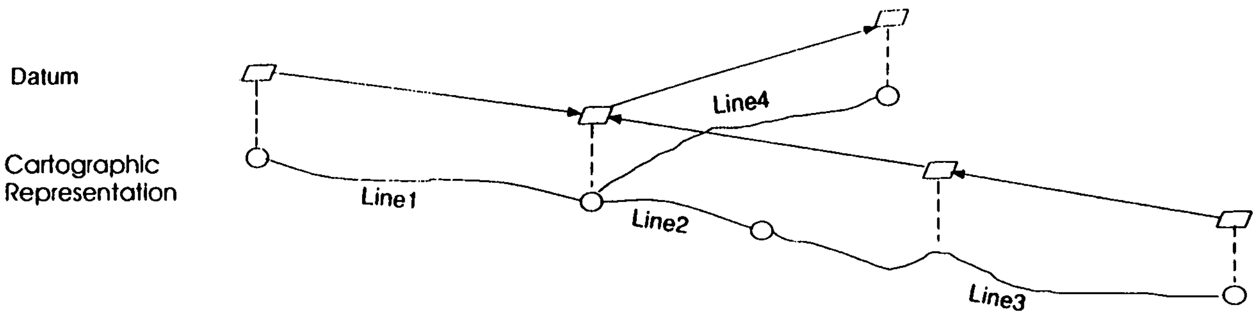

Figure 4. Datum and cartographic representation.

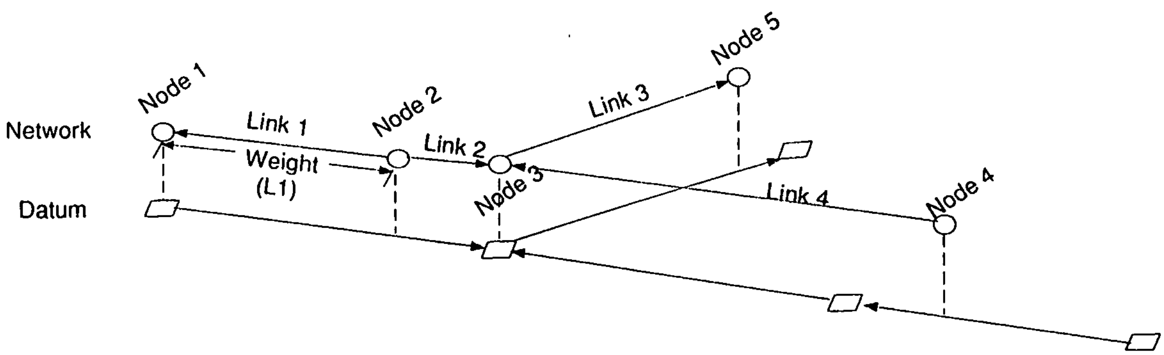

Figure 5. Datum and network.

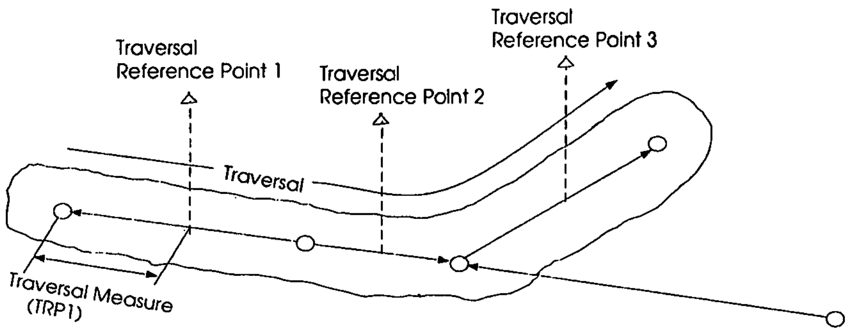

Figure 6. Traversal and traversal reference points.

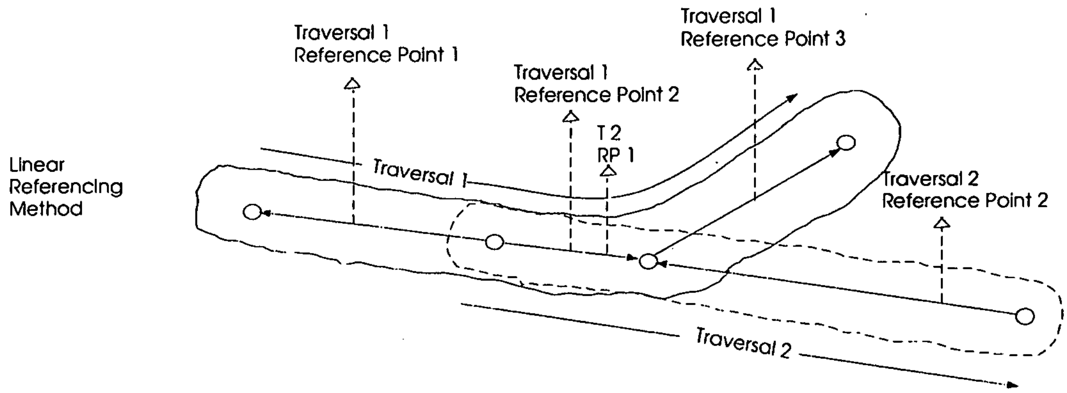

Figure 7. A linear referencing method can have many traversals (many traversals/many links).

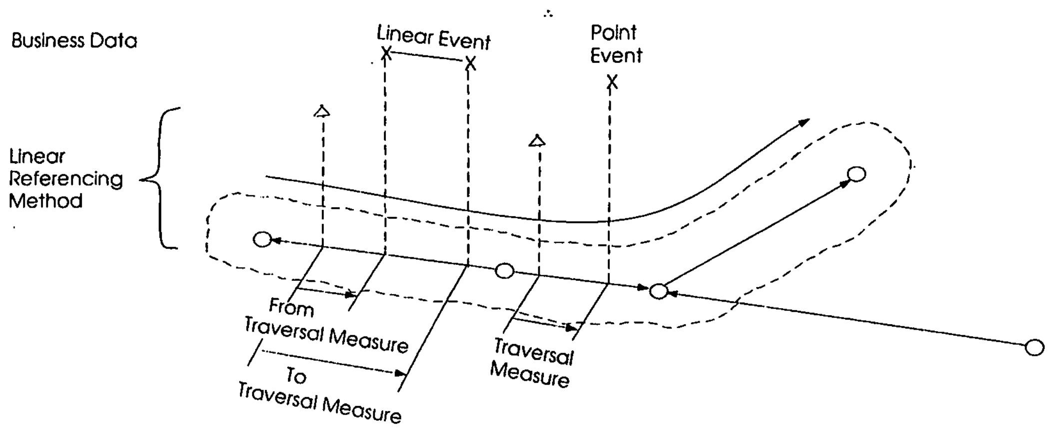

Figure 8. Linear referencing method with business data.

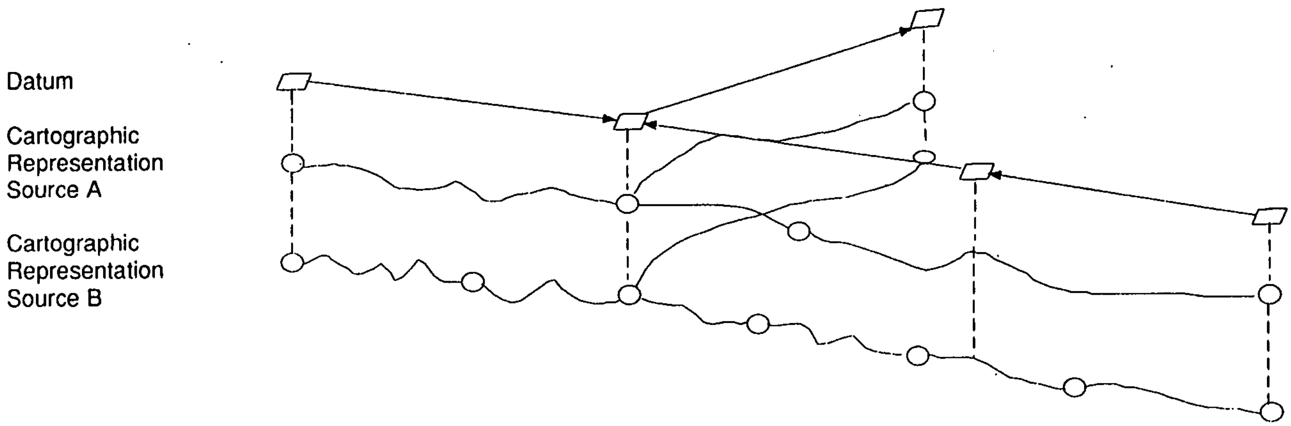

Figure 9. A datum can have many cartographic representations.

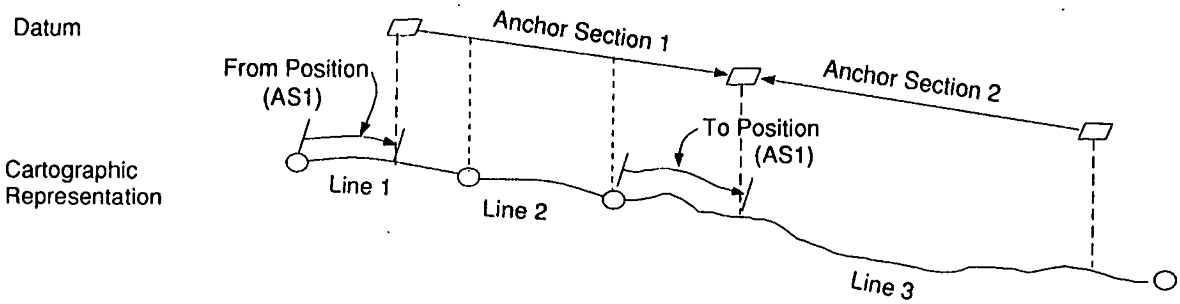

Figure 10. Many anchor sections/many lines.

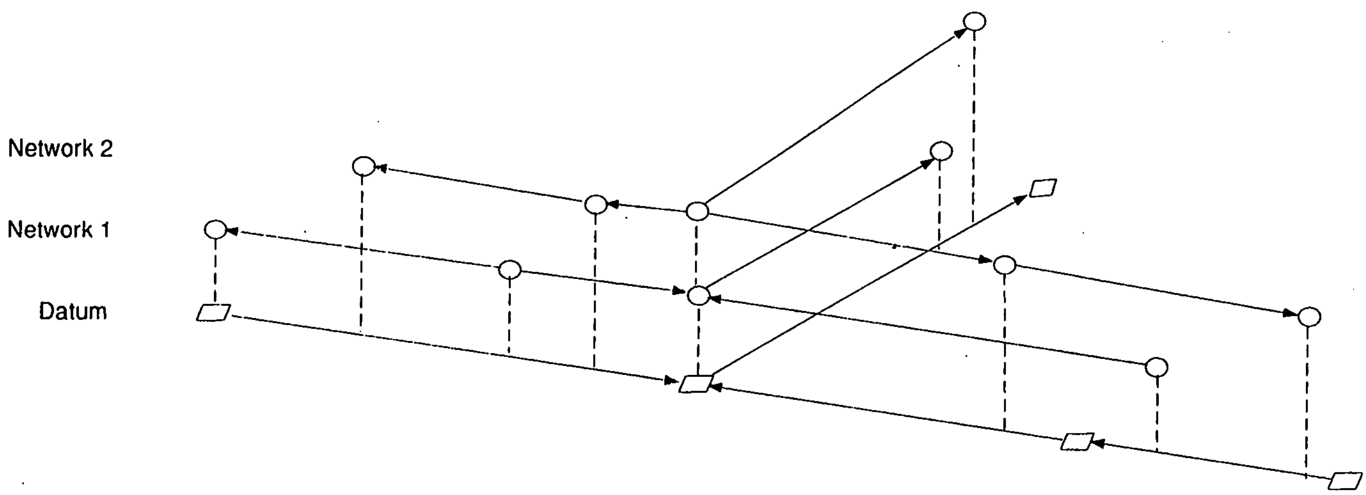

Figure 11. Many networks can be referenced to a datum.

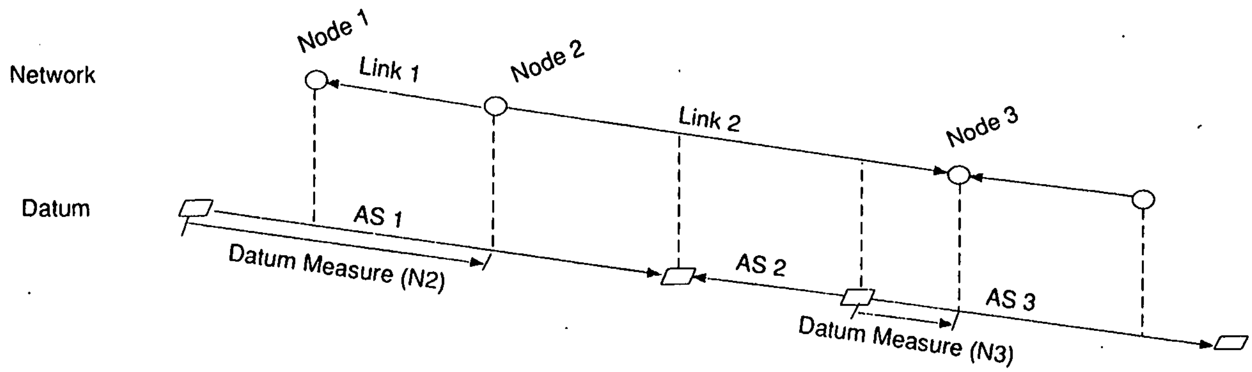

Figure 12. The anchor section/node association.

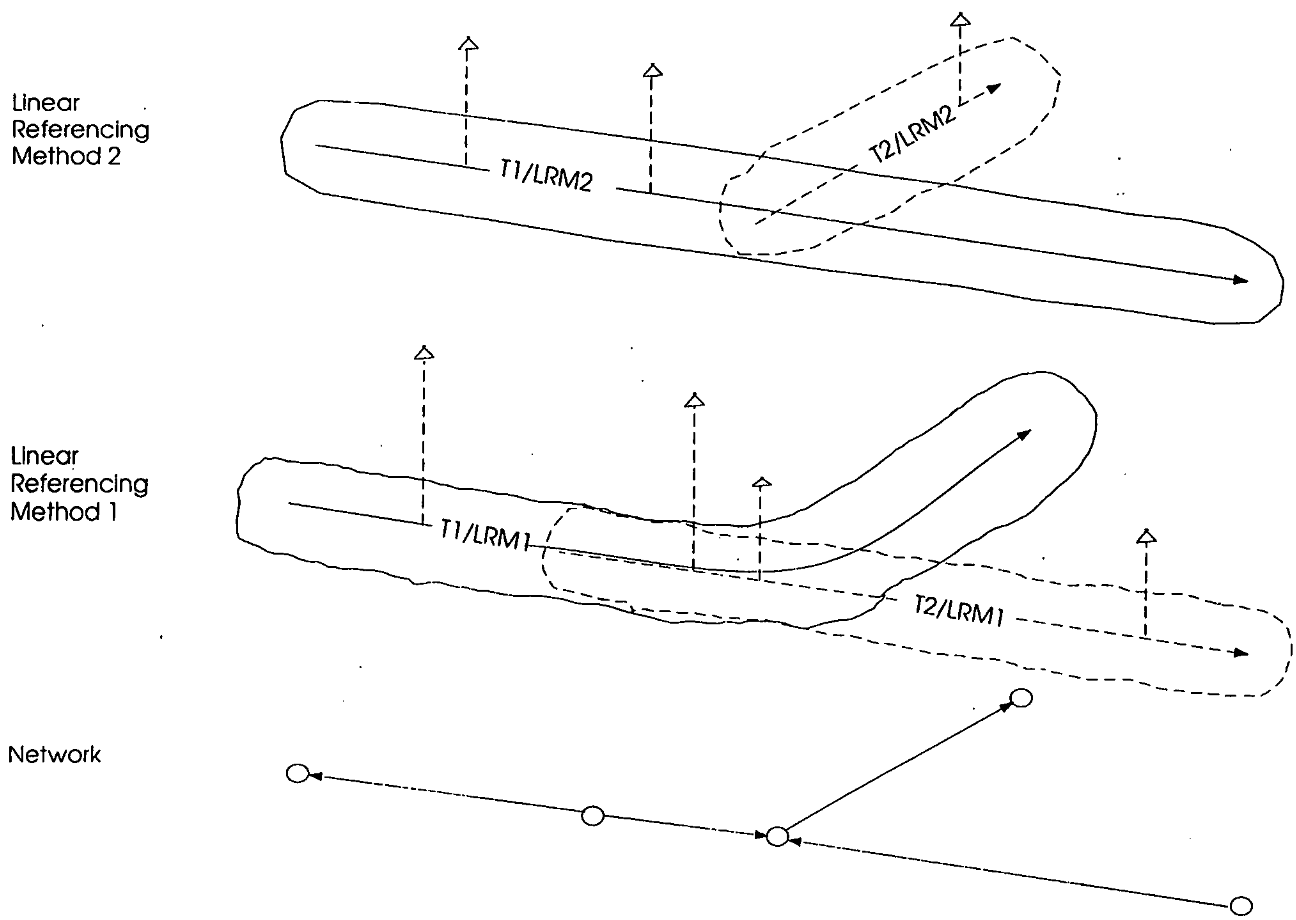

Figure 13. A network can have many linear referencing methods.

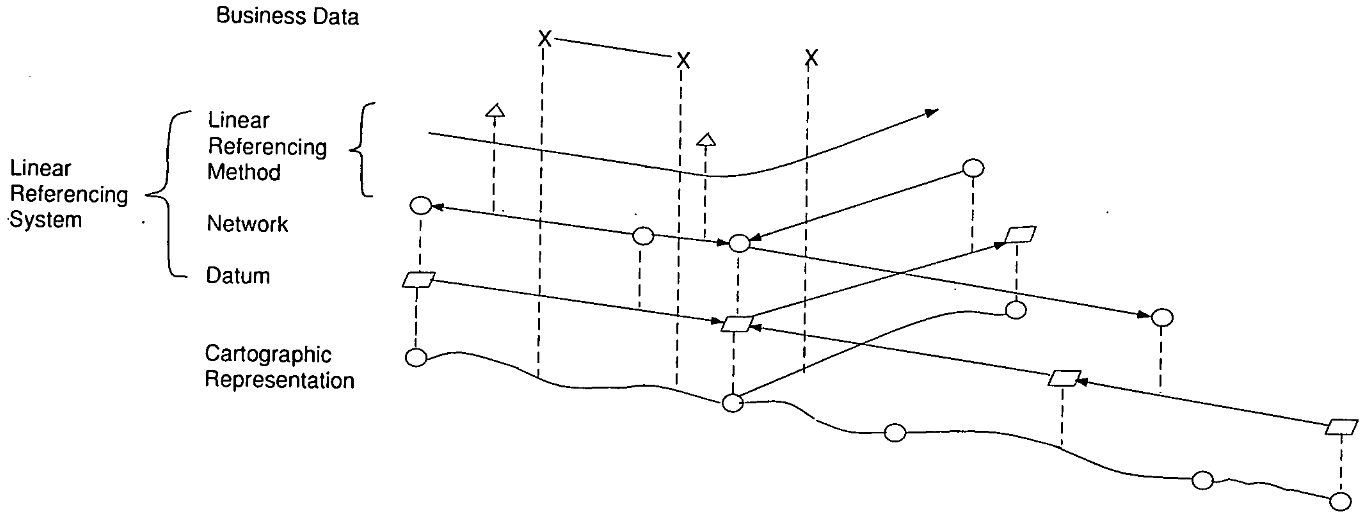

Figure 14. Linear referencing system with business data and cartographic representation.

TRANSPORTATION RESEARCH BOARD
National Research Council
2101 Constitution Avenue, N.W.
Washington, DC 20418