A Comprehensive Process for Linear Referencing

Kevin M. Curtin, Greta Nicoara, and Rumana Reaz Arifin

Abstract: *This paper identifies and analyzes linear referencing as a spatial process. This is a significant departure from the existing research that defines linear referencing as a set of objects. The critical issues in the development of such a process are identified, including the determination of network representations and topology, the route structure, the determination of measures along linear features, the creation of event data, and the display and analysis of those events. A model is presented that delineates this process for streamlining the implementation of linear referencing, and providing a structure that can manage the increasing level of complexity in spatial data.*

Keywords: linear referencing, GIS-T, transportation, dynamic segmentation

INTRODUCTION

The primary objective of this paper is to identify, demonstrate, and analyze the necessary and sufficient requirements for exploiting linear referencing. A linear referencing process is developed and presented that expands on the extant linear referencing data models, methods, and systems that have appeared in the literature. This process is intended to provide a framework for the implementation of linear referencing among an expanding group of Geographic Information System (GIS) users.

Linear referencing can be used by many organizations, industries, and institutions that work with linear features, such as road-management organizations, transit organizations, oil and gas exploration industries, and water-resources managers, to name only a few. The common element among these industries is their use of linear features and the need to reference a position or measure along those features.

As GIS becomes more prevalent among an increasingly diverse and rapidly growing set of users, including small to midsize municipalities, government agencies at every level, and private businesses, there is an increasing demand for more sophisticated approaches to data management. When the network databases that have long been modeled in GIS (Curtin 2007) are an important element of the analyses undertaken by these groups, the need to successfully implement linear referencing becomes an important issue, and a process for linear referencing is essential.

This research presents a comprehensive process for linear referencing. In the following section, linear referencing is formally defined and its advantages outlined. This is followed by a comprehensive literature review that discusses both the applied use of linear referencing—particularly in GIS—and the theoretical models and methods that have been developed. Based on this review, a seven-step process for linear referencing is presented and its use is demonstrated through a case study of the city of Richardson, Texas. Conclusions regarding the potential use of this process and opportunities for future research are discussed.

Linear Referencing Defined

The term *linear referencing* emerged from engineering applications where it was preferable to locate a point along a linear feature (often roads) by referencing that location to some other well-defined location, rather than using classical geographic coordinate systems. The most familiar illustration of linear referencing is the mile markers along U.S. highways (Federal Highway Administration 2001, Federal Transit Administration 2003).

Determining locations with linear referencing differs from traditional geographic coordinate and reference systems (latitude-longitude, Universal Transverse Mercator (UTM), state plane, etc.) for the underlying entity used as a basis for measurement is not the earth, but is rather a linear feature or a set of linear features organized into a network. Just as there are myriad coordinate systems for the globe, there are multiple linear referencing systems. A common definition for a linear referencing system (LRS) is a support system for the storage and maintenance of information on events that occur along (or within) a transportation network. In this context, an LRS consists of an underlying transportation network that supplies the geographic backbone for the location of events, a set of objects with well-defined geographic locations (also known as a datum), one or more linear referencing methods (LRMs), and a set—or sets—of points or linear events that should be referenced to the underlying network. This paper will demonstrate that in many cases the underlying network and the datum are one and the same.

An LRM can be defined as a mechanism for finding and stating the location of an unknown point along a network by referencing it to a known point (Vonderohe, Chou et al. 1997). More specifically, an LRM is a process for determining a previously unknown location based on (1) a defined path along the underlying transportation network, (2) a distance along that path measured from a known datum location, and (3) optionally an offset from the path. There are several different common types of LRMs that differ based on the parts of the network used for referencing and the ways in which measures and offsets are calculated (Nyerges 1990).

Applications and Benefits of Linear Referencing

Linear referencing can be applied to any network-based phenomenon. Given the historical development of the technique of linear referencing, however, transportation applications dominate the literature. Some of the more common transportation uses are the mapping of accident, traffic stop, or other incident locations, and asset-management functions such as the recording of pavement conditions or the location of street signs, streetlights, bridges, or other traffic-related objects. Despite the historic concentration on transportation applications, significant benefits result from using linear referencing for applications in many different fields. In hydrologic modeling, linear referencing can be used to locate flow gauges along rivers or monitoring stations along creeks or pipelines. In utility facilities management, linear referencing can be used to model and display the attributes of the distribution network.

For *any* network application, using linear referencing has several primary benefits. First, locations specified with linear referencing can be readily recovered in the field and are generally more intuitive than locations specified with traditional coordinates. Secondly, linear referencing removes the requirement of a highly segmented linear network based on differences in attribute values. More specifically, many network attributes do not begin, end, or change values at the same points where the network is segmented: *i.e.,* speed limits do not always change at intersections, pavement quality can change at any point along a road, and stream widths can change at many different points along a stream channel. If the changes in the values of all network attributes were used to segment the network so that each segment could have a unique attribute value, this would result in an increasingly segmented (and therefore larger) database. The implementation of linear referencing allows an organization to maintain a network database with many different attribute events associated with a single, reasonably small set of network features. The implementation of linear referencing thus reduces the redundancy and potential error within the database, and it facilitates multiple cartographic representations of network attribute data.

Literature review

The literature pertinent to this research falls broadly into two areas: the theoretical data models for linear referencing that have been developed and the implementation of linear referencing in GIS.

THEORETICAL LINEAR REFERENCING DATA MODELS

Those who wish to apply the principles of linear referencing within GIS face a daunting set of theoretical linear referencing data models. Perhaps most significant among these are the models developed under the auspices of the National Cooperative Highway Research Program (NCHRP) project 20-27, which developed a succession of linear referencing data models in consultation with a wide range of academicians, practitioners, and transportation policy makers (Vonderohe, Chou et al. 1997; Vonderohe, Adams et al. 1998; Koncz and Adams 2002; Koncz and Adams 2002; Koncz and Adams 2002). These efforts concentrated on identifying the most basic underlying elements in linear referencing systems and methods to provide a generic data model. To eliminate known difficulties stemming from differences in terminology (Dueker and Vrana 1992), these researchers comprehensively defined terms, concepts, and relationships that could apply across application areas and geographic scales of operation. Although these models comprehensively define many objects and relationships, "a literal interpretation of the NCHRP model would be too difficult to . . . implement" (Scarponcini 2001).

While some have concluded that a single unified linear referencing system could meet the needs of all transportation users (Fletcher, Expinoza et al. 1998), much of the NCHRP and other linear referencing modeling work has focused on the decoupling of topological, graphical, positional, and attribute characteristics of transportation objects to facilitate data sharing within enterprises (Kiel, Pollack et al. 1998; Dueker and Butler 2000) and the translation of locations between linear referencing methods (Scarponcini 2002). Another model suggests that the attributes that would traditionally be referenced to the network can be the primary object to be stored in the spatial database, while the location and shape information is encapsulated with the attribute (Sutton and Wyman 2000). Lastly, other efforts have concentrated on identifying essential data models that allow for flexible definitions of (and relationships between) transportation database objects (Curtin, Noronha et al. 2001), including those objects related to linear referencing.

Linear Referencing and GIS

The vector data model that has dominated the application of geographic information science (GIScience) since the inception of the discipline is widely recognized as an extraordinarily useful data structure for transportation systems and other network processes (Curtin 2007). The ability to reference events to features in that data structure has long been identified as an essential functionality within GIS (Nyerges 1990). Several GIS software packages currently offer tools to assist in the generation of spatial features and events for the purpose of linear referencing (Goodman 2001). The documentation for such tools, however, is focused primarily on how to create the events and define the measures within their systems (ESRI 2001, ESRI 2003). There is very little—if any—insight into how these events and measures should be captured, analyzed, or maintained. It is difficult for users to implement linear referencing without a well-defined process to follow.

Historically, the practitioners of linear referencing in large transportation agencies knew that distance measurements collected in the field (sometimes using measuring wheels or other highly accurate distance measuring tools), and stored in tabular form (outside of the GIS), were superior to the digital data representations that—for decades—suffered from a persistent lack of positional and attribute accuracy. With the proliferation of GIS over the past decade, smaller users, including government agencies and private businesses, are capable of generating the most accurate geospatial data available for their areas. Often this is more detailed than commercially or nationally available products. These smaller users often store their data in a single location, using a single spatial data format, and they maintain that data with a small staff. This spatial representation is accepted as the highest quality representation of their network. The GIS digital data is the datum; the well-known street intersections or other captured points stored in the GIS are the anchor points for linear referencing. Given this acceptance of the data stored in the GIS, there is no need for an artificial separation of the cartographic representation and the analytic network database, which has been one of the foundations of linear referencing modeling efforts.

The use of linear referencing is a way to improve the return on the investment made in adopting geospatial technologies. When a street centerline geodatabase is being captured and maintained, building an LRS is a logical step forward that expands the number and diversity of applications that can be implemented (Noronha and Church 2002). As the paucity of literature regarding the use of linear referencing suggests, however, this valuable tool is infrequently implemented by users other than major transportation agencies. The question this research seeks to address is why do these GIS departments not implement linear referencing when the tools for doing so are readily available? The authors believe that the answer lies in the absence of a clearly defined process for implementing and using linear referencing within GIS. The following section outlines such a process.

Linear Referencing as a Process

This paper presents an iterative, seven-step linear referencing process (shown in Figure 1). The first step of this process is to identify an application to which linear referencing is pertinent and to use that information to decide what network representation should be employed and the topological rules that must be followed. The second step determines the route structure—or the underlying datum—to which events can be linearly referenced. The third step identifies the way in which measurements will be made along those routes and the fourth step defines the way in which linear events will be defined, captured, and maintained. The fifth step concerns the cartographic output of linearly referenced events, and the sixth step outlines ways of analyzing those events once they have been fully referenced. The final step in the Linear Referencing Process (LRP) is to maintain the linearly referenced data in such a way that it can be shared with other agencies, used for many different applications, and queried based on historical conditions. Each of these seven steps are outlined below and tested using a case study of the city of Richardson, Texas, a midsize city with a population of approximately 100,000 in the Dallas/Fort Worth metropolitan area.

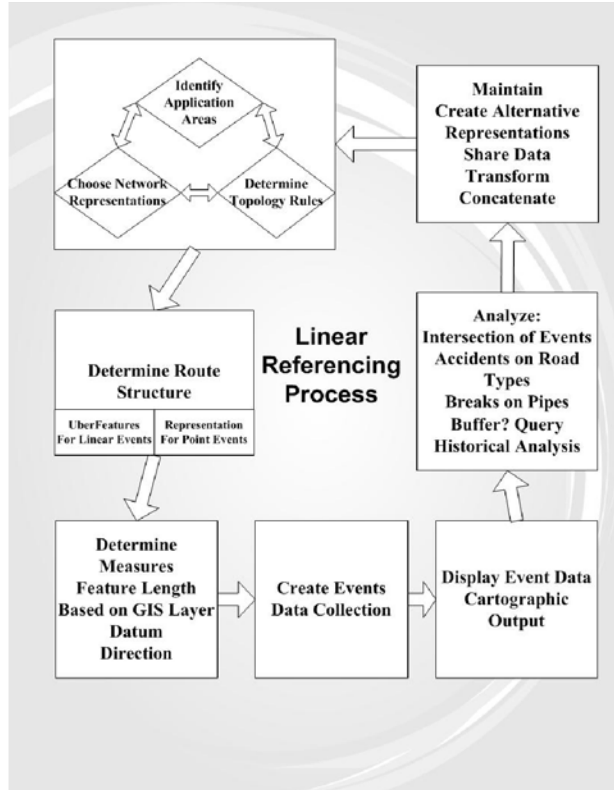

Figure 1. An iterative seven-step linear referencing process

Determine Application, Network Representation, and Topology

As described previously, myriad applications can benefit from the implementation of linear referencing. Although it would be ideal if all applications could rely on the same measurement techniques, use the same network databases, and employ the same types of network analyses, this is simply not the case. For example, road networks and municipal water networks are fundamentally different in many ways. Flow in road networks concerns independent mobile entities (cars, trucks, bicycles, etc.), while flow of water through a network of pipes is determined by demand, pressure, and elevation among other factors. Similarly, fundamental differences exist in the analytical and cartographic needs for electrical networks, gas or oil pipeline networks, or river networks. Although they all depend on a network structure, the attributes and the analytical methods associated with these different network types require different linear referencing specifications.

Therefore, the first step in a linear referencing process is to define which network datasets (and what representations of those networks) are to be employed for the application at hand. In some

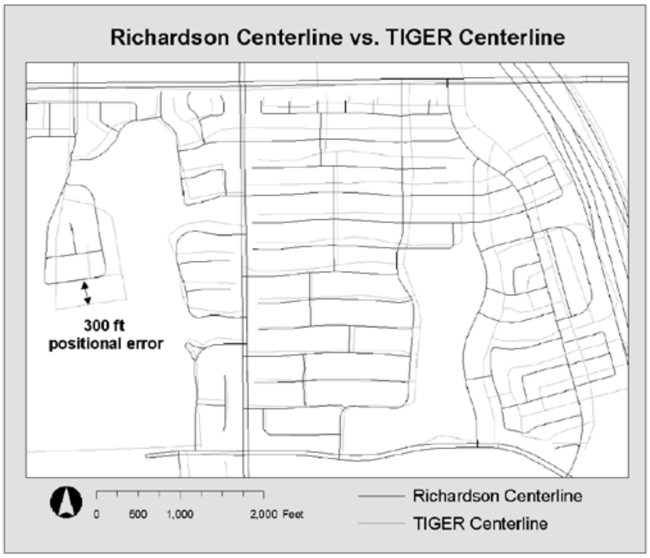

Figure 2. TIGER streets and corrected Richardson streets

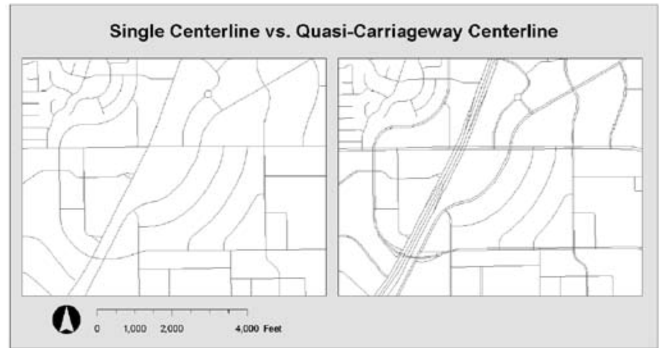

Figure 3. Simple centerline over quasi-carriageway

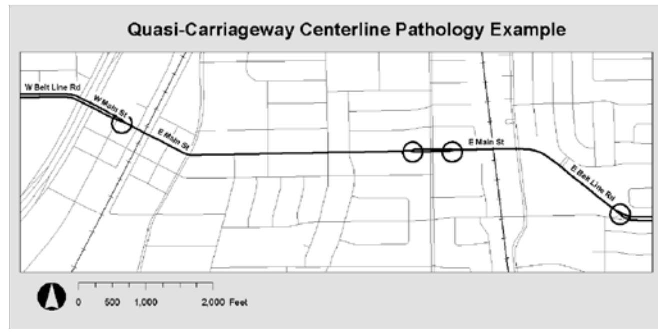

Figure 4. Quasi-carriageway pathology example

cases, several different representations are available for the same network, and these competing representations may differ based on their source, their coverage, their attributes, or their topological structure. Moreover, several network datasets may need to be used together, such as employing both the water distribution network and the road network for the emergency response to an incident such as a water main break.

As an example of determining application, network representation, and topology for linear referencing, we turn to the case study area—the city of Richardson—to examine a street centerline–based application for pavement management. In the city of Richardson, the road centerline was originally generated from corrected Topologically Integrated Geographic Encoding and Referencing (TIGER) data that had been provided by the North Central Texas Council of Governments (NCTCOG). It has since been modified significantly; most recent changes involve the use of six-inch orthophotos, a practice becoming more common among municipalities as high-resolution imagery is becoming increasingly available. As such, the GIS road network representation is generally trusted to be the most accurate positional reference data within the municipality. In terms of linear referencing, this means that the road network itself can serve as the datum, unlike in most theoretical linear referencing data models that require a separate datum to compensate for the inaccuracy of the centerline dataset (see Figure 2).

One common application for the street centerline database is the inspection and maintenance of the pavement on those streets. In Richardson, the street department is provided with a cartographic representation of the street centerline file, which it uses to identify pavement surface types (concrete, asphalt, or concrete with asphalt overlay). These attributes are then input into the GIS as linearly referenced events for the purpose of querying the street centerline database to determine the amount and age of each pavement type, and subsequently to support street maintenance projects.

Given that an application has been chosen—in this case, pavement management—the specific street centerlines representation must be chosen. There can be significant differences among representations of the same road network, even within the same municipality or agency. In this paper, we discuss two primary network representations: a simple centerline representation and a quasi-carriageway representation. A simple centerline representation reflects the most common road network structure: a fully planar TIGER-based representation with centerline splits at all feature intersections, regardless of the nature of the actual street intersection at that point. Given the history of this network structure, it is the most commonly found across applications and agencies, although nonplanar alternatives exist, and several researchers have recognized the increasing demand for GIS-T data models that support lane-level operations rather than forcing all applications to conform to centerline or carriageway representations (Fohl, Curtin et al. 1996; Miller and Shaw 2001). The single centerline representation is generally used for generating road length measurements and for cartographic applications. In contrast, a quasi-carriageway centerline representation is a more detailed version of the network with individual features representing the flow of travel along major roads, particularly those that are divided by central medians (as shown in Figure 3).

Generally, medians only separate major city roads such as arterials and collectors, therefore most residential streets have a single centerline representation. In some instances, however, roads may alternate from a single to a double representation because of the lack of or presence of a median (thus the "quasi" for not all streets are represented with two features showing the different directions of travel). Such network pathologies can cause difficulties in the determination of what constitutes a route in the linear referencing system (see Figure 4).

Although Richardson does not have any one-way streets, the carriageway or quasi-carriageway representation may be more appropriate for areas with large numbers of one-way streets to maintain information about traffic-flow direction. However, Richardson also maintains traffic-flow direction attributes, represented by bidirectional flow in single line segments and flow as going either with or against the digitized direction where double lines are used for street segments. Additionally, street attributes are appended directly to the centerline segments (such as block ranges for address geocoding, speed limits, pavement types, and rights-of-way, among others).

Officials in Richardson determined that it was imperative that the database not be segmented any more than necessary to reduce the network database size and to encourage accurate routing across the network. Therefore the topology of the quasi-carriageway representation had to differ from the fully planar representation, so that no segmentation of features should occur unless there is a physical intersection of streets where traffic can flow from one street to another. In cases where there are bridges, tunnels, overpasses, or other split-grade intersections, there should be no split of the features.

To summarize, for the case study of Richardson, the first step in implementing linear referencing was to choose an application (pavement management), to determine the best network representation (quasi-carriageway), and to determine the appropriate topology (nonplanar at grade-separated intersections).

Determining Route Structure

The next step—and perhaps the most challenging—in the process of linear referencing is the determination of the route structure. In this research we define a route as the largest individual feature that can be uniquely identified and to which events can be linearly referenced. This definition differs from the common notion of a route such as a bus route that may traverse several or many different features along an established course of travel. Additionally, this definition of route is mirrored most closely by the NCHRP 20-27 definition of an "Anchor Section." Features such as roads, railroads, creeks, and, ultimately, any linear feature can become the underlying element of a route.

Although routes could be created from individual street segments in the network database, this would eliminate one of the primary benefits of linear referencing, that an event spanning many street segments can be maintained as a single object in the database. To avoid this, the composite set of road segments that will constitute a route should be longer than the events to be referenced. For example, the pavement type of a street may often be the same for the full extent of a road within the city limits. Ideally, all the segments that make up that road should constitute a single route, so that one event can define the pavement type for that route. Thus, the determination of what constitutes a route must be made with regard to the largest event along that route to minimize unnecessary event segmentation.

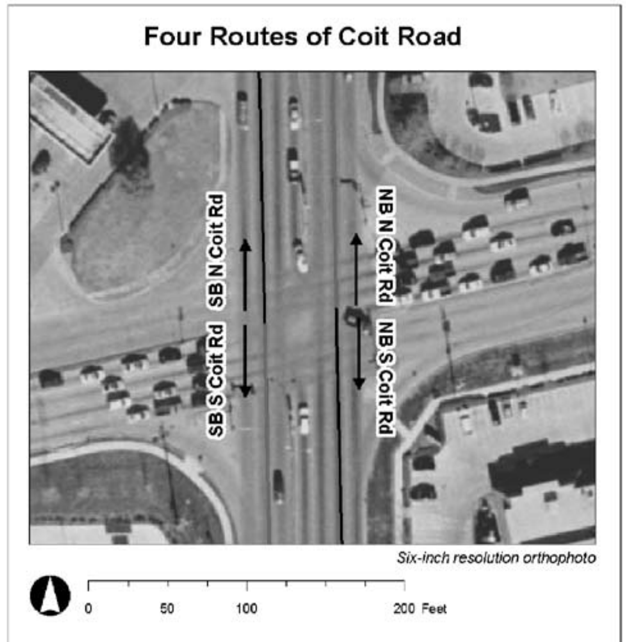

Figure 5. Route structure example

The use of an entire road spanning the geographic study area is the approach that is widely used in statewide linear referencing systems (Texas Department of Transportation 2003). However, when a single road feature has two or more names associated with it, multiple routes could be established based on the street names. Similarly, if applications (such as address geocoding) will depend on directional prefixes on the roads, it may be desirable to have routes that correspond to both street name and direction. Moreover, if a carriageway or quasi-carriageway network representation had been chosen in step one of the linear referencing process, then each of the carriageways may be a separate route, even if the street name and direction prefix are the same.

In the Richardson case study, the underlying street network database has a quasi-carriageway representation. Topologically, splits on this database are enforced only where at-grade intersections occur, where road names change, or where address block numbers change.

Given these representational choices, the routes for linear referencing have been based on the street name with directional identifiers. When carriageways (double lines) exist, a separate route is established for each direction of travel (*e.g.,* northbound and southbound). Because residential roads are all single lines, each full street name/direction prefix constitutes one route. When a street has no directional prefix, the entire street becomes a route. For example, the two sets of segments that make up Coit Road were actually split into four separate routes according to the following route-name attribute: NB N Coit Rd, SB N Coit Rd, NB S Coit Rd, and SB S Coit Rd (see Figure 5).

In summary, the route structure determination decision depends on all the decisions made in the prior step of the linear referencing process in addition to decisions regarding what attributes are to be linearly referenced. This necessitates an evaluation of the trade-off between the level of detail that one wants to reference against and the concomitant increase in the number and segmentation of routes.

Determining Measures

Once the route structure is determined, the third step in the linear referencing process is to determine measures along those routes. There are three primary considerations when setting measures along routes: (1) the unit of measure that is most appropriate, (2) the source for the measure values, and (3) the direction of increasing measure values. The most appropriate unit for measures along routes is a function of the application for which the linearly referenced features will be employed. Because unit conversions are a commonplace function in GIS software, changes in units require only that both the measures along the routes and the measure information associated with the event data remain consistent. However, if more than one unit is required for the same routes, this may require that a second set of routes be maintained. For some applications, the measure along a feature may be given as a percentage of the distance along the feature rather than as an absolute distance.

As discussed previously, the source data for measure values is a subject of substantial debate. Historically, data collected in the field and stored in databases external to the GIS were considered substantially higher in quality in terms of spatial accuracy than the digital spatial data employed by the GIS. This fact led linear referencing researchers to develop well-defined objects (such as anchor points and anchor sections in the NCHRP 20-27 model) that were based on the well-defined location of points and segments collected in the field. Measures were computed along the network based on their distances from these objects. These measures could be associated with cartographic representations of the network, but differed from the distance values computed internally by the GIS. Today, the digital data landscape has completely changed in this respect. The data maintained in GIS departments is almost universally considered the highest-quality spatial data available, and any improvements or corrections in feature locations are almost immediately transferred to the GIS database. Additionally, consortia of municipalities and nongovernmental organizations collaborate extensively on data collection and maintenance. Therefore, those implementing linear referencing today are free to base their measures on the positions and lengths of features as they are computed and stored within the GIS. There is no need to maintain a separate set of features from which measures are computed. Once again, the GIS data *is* the datum.

The third consideration when determining measure values is the direction of increase of those values. Once again this depends on the previous steps in the linear referencing process. Generally speaking, the direction of increase should be consistent with the needs of the application chosen in the first step of the process, and it should be logically consistent with both the topological network design and the route structure determined in step two of the process. For example, consider an application concerned with a river network and multiple routes associated with different branches of the network. Because activities along the river network are likely to be associated with the flow of the river, the measures along the routes should be consistently associated with either the upstream or the downstream direction of the river network. Similarly, if routes are designed to conform to the direction of travel along a street network, then increasing measures may most appropriately conform to those directions. However, if applications associated with addresses are of primary interest, then the direction of increasing measures should likely conform to increasing address ranges.

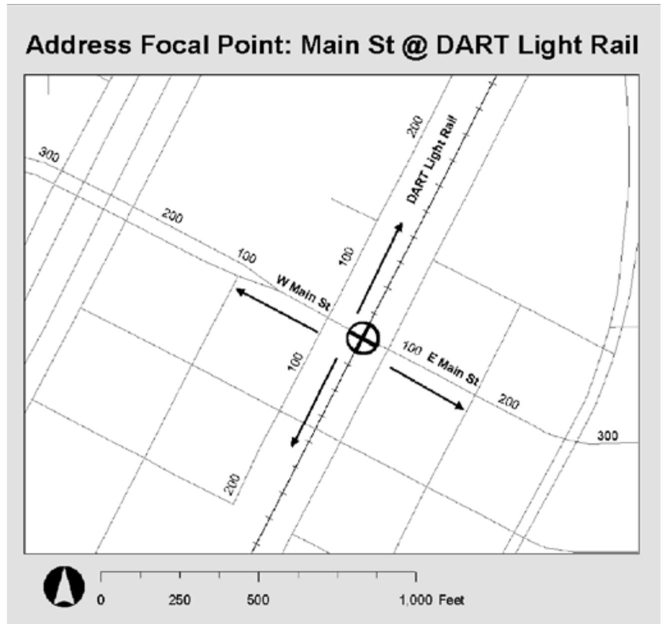

Figure 6. Measures increasing from the focal point

Turning to the case study area, in the city of Richardson, measure values were based on feature length because the GIS data is accepted as the datum and the geometric lengths of the segments are accepted as sufficiently accurate for most city needs. In Richardson, route measure units were based on the projection of the data (state plane) that customarily employs feet as the unit of measure. For Richardson, route measures were designed to increase in accordance with the addressing scheme for the city. More specifically, measures increased from the origin point of the city's street grid that is the intersection of Main Street and the Dallas Area Rapid Transit (DART) rail station located in the historic downtown section (shown in Figure 6). This is the same point used as the dividing point for routes.

Create Events

When the first three steps of the linear referencing process have been completed, a set of routes have been built from the underlying network for the chosen application. These routes are informed

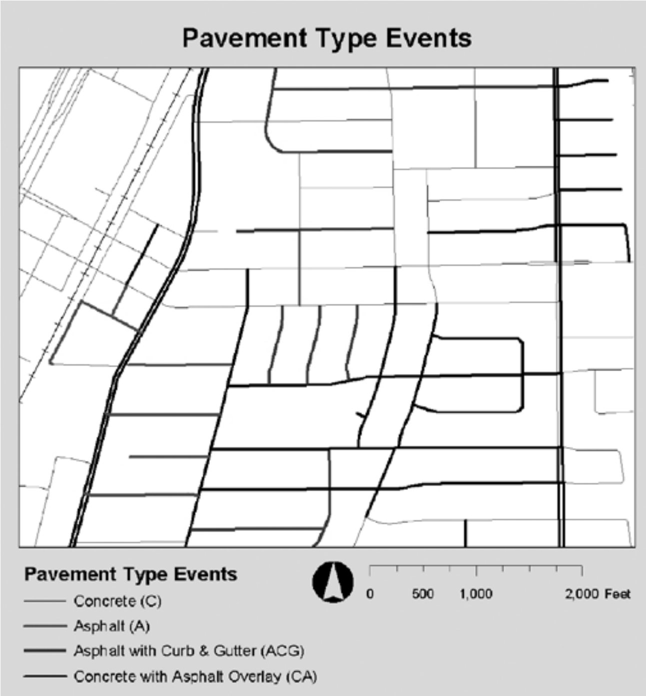

Figure 7. Pavement types represented as linear events

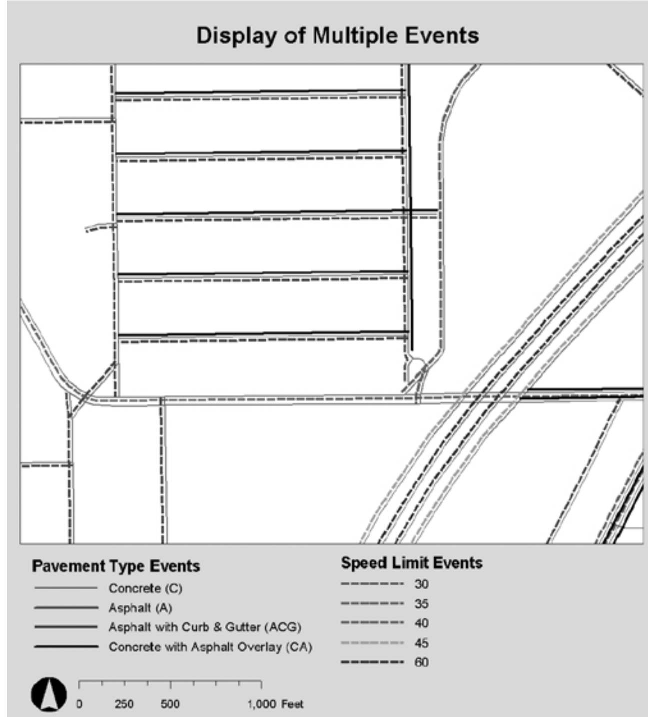

Figure 8. Cartographic display of related events

with measure information based on application needs and topological structure. The next step in the process is the collection of event data to be referenced to the newly created routes. For many users, this step is what first comes to mind when considering linear referencing. However, without having completed the first three steps of the process as outlined previously, a great deal of effort in event data collection could be wasted.

Event data are occurrences along the network. Events can be point or linear in character. Point events represent some object at a specific measure along a linear feature. Examples include traffic accident locations or traffic control devices such as signs or signals. Linear events often correspond to objects that have a consistent attribute along the network. Examples include pavement type or condition, speed limits, traffic volumes, or pipe widths along a water network. An event is known as a "traversal" in the NCHRP 20-27 efforts.

Event data can be collected or created in many different ways. Events can be digitized from a range of cartographic products including both paper maps and aerial or satellite photographs. Events can be collected in the field either by direct observation of attributes or locations by personnel or with GPS receivers that capture locations for subsequent input to the GIS as events. Custom software tools exist to facilitate event data collection, conversion, and maintenance. Existing point data can also be converted to point events, or event tables can be populated manually based on known locations and attributes. As needed, event data can be exported as a set of stand-alone features; otherwise, it can remain in tabular format. Perhaps most important, whichever method of event data collection is chosen, the events must be structured in such a way that they can logically be associated with the routes determined in the second step of the process.

In the case of Richardson, the initial objective for implementing linear referencing was to more accurately represent surface pavement type. Data on the locations of different pavement types had been collected in the field and recorded on a paper map. This data had been transferred to the network database, but was represented inaccurately for road segments had not been split for this purpose. Therefore, the pavement type attributes had to be approximated to the nearest intersection. With the implementation of linear referencing, more precise pavement type events could be associated with the previously determined routes without further splitting the network features solely for the purpose of attribute differentiation (shown in Figure 7).

Display Event Data; Cartographic Output

One of the most powerful arguments for the implementation of linear referencing within a municipality is the ability to accurately display the event information for in-the-field use by employees, or for higher level analysis and decision making. The ability to display multiple attributes associated with networks provides both opportunities and challenges. The increased information available for display can provide new insight into the problems under examination, but can also lead to poor cartography because of graphical clutter and information overload. Therefore, the next step in the process of linear referencing is to carefully choose the parameters for display of the linearly referenced information.

The decisions regarding display of event data depend on several factors, including the media on which the data will be

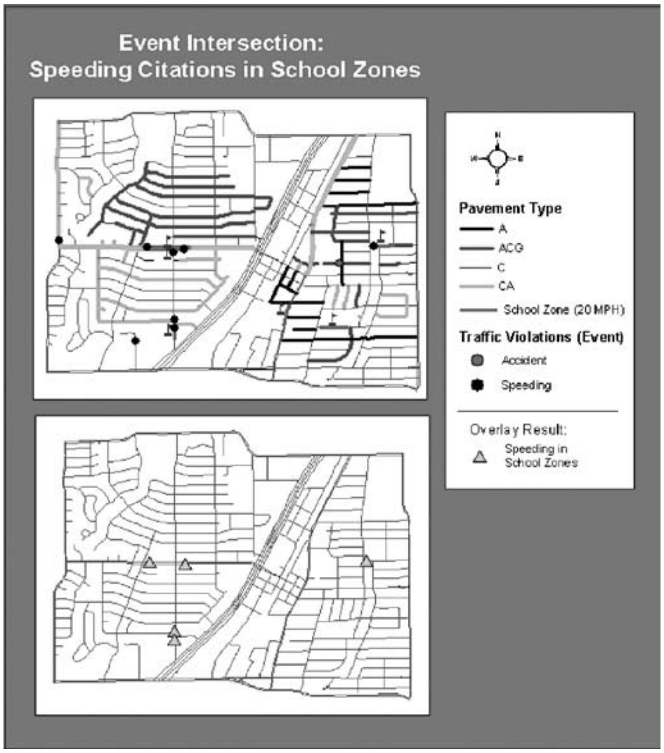

Figure 9. Intersection of linear and point events

displayed, the scale (or scales) at which the data will be displayed, and the computation of event locations within the GIS. Their subsequent display is a process that is often referred to as dynamic segmentation. One visual benefit of dynamically displaying event data is the ability to display multiple linear events along the same feature, using varying event offsets. Common examples of this practice include subway maps and bus route maps. Richardson used the pavement type events and speed limit events to create a single cartographic display of the locations of these two attributes more clearly than could have been accomplished using other cartographic procedures (see Figure 8).

Analysis with Linear Referencing

The ultimate goal of implementing linear referencing (or any other process) in a GIS is to increase the ability to perform a diverse set of analyses. With routes and event data in hand, analysis can then be performed on the event data, through techniques such as overlays, intersections, and other techniques that are part of the linear referencing capabilities of most GIS software. For example, in Richardson, the association between traffic violations and school zones was explored through an intersection of speed limit linear events (specifically those events associated with school zones) and traffic violation point events (as shown in Figure 9).

Linear referencing allows an entirely new set of database queries to be made that differs from queries based on the underlying network. For instance, the storage of data as event tables enables historical queries if events are date-stamped. However, while significant analytic capability is added through the linear referencing process, other traditional GIS analytic capabilities are lost. One example is the loss of traditional road network functions, such as shortest path determination or routing, because of the loss of nonplanar topology that results when the segments that comprise the routes must be merged to satisfy the need of a large-enough route to minimize event differentiation.

Table 1 contains a comprehensive list of common GIS analytical tools and describes whether or not these tools can be used with linear referenced event data. It is important to note that the tools for network analysis, geostatistical analysis, and geocoding currently cannot be used with the linearly referenced events. These are functions that are universally used for network-based applications, and their extension for use with linearly referenced data would represent a substantial advance for network analysis.

DATA MAINTENANCE

To keep the newly created linear referencing system functional, it is important that the route and event data be maintained properly. As changes are made to the original road file, the same changes must also be reflected in the routes. There are extant tools for setting appropriate topology rules, which, in turn, enable multiple route feature classes using different linear referencing methods to be built and maintained on a common reference layer of roads. If the original roads and all associated routes are united in a topology, this also enables the simultaneous editing of these multiple feature classes.

Furthermore, measure values need to be maintained if roads are ever removed or have their course altered over time. Or if even more precise measure data becomes available, routes can be calibrated to reflect this new data. While the maintenance of the linearly referenced data and the underlying network may not be the most fascinating of tasks, it is necessary to keep an implementation of linear referencing functioning.

CONCLUSIONS

Although there are a multiplicity of high-quality theoretical models and methods of linear referencing, and a substantial number of tools for implementing linear referencing in a GIS context, there has until now been no explicit process for implementing linear referencing. This paper presents a comprehensive process for linear referencing that consists of outlining the application for which linear referencing is intended, defining the nature of the underlying network, identifying the underlying routes, specifying a system to measure locations along those routes, collecting and storing event data, performing analysis with those events, and maintaining the linear referencing system. This seven-step linear referencing process is intended as the basis from which any application of linear referencing can proceed. It is hoped that this structure will allow the extraordinarily useful set of linear referencing tools to be more widely accepted among GIS users, and will encourage GIScientists to more closely examine the processes behind these tools and thus increase the ability to perform robust geospatial analyses.

Table 1. Event Analysis with Different Tools

Tools	Event Analysis
1. Analysis Tools	
❊ Extract	
o Clip	Yes
o Select	Yes
o Table Select	Yes
❊ Overlay	
o Intersect	Yes
o Union	No
❊ Proximity	
o Buffer	Yes
❊ Statistics	
o Summary statistics	No
2. Conversion Tools	
❊ From Raster	No
❊ To dBASE	No
❊ Geodatabase	
o Feature Class to Geodatabase	Yes
o Table to Geodatabase	Yes
❊ To Raster	No
3. Spatial Statistics Tools	
❊ Analyzing Pattern	
o Average Nearest Neighborhood	Yes
o High-Low Clustering	Yes
o Spatial Autocorrelation	Yes
❊ Mapping Custers	
o Cluster and Outlier Analysis	Yes
o Hot-Spot Analysis	Yes
❊ Measuring Geographic Distribution	
o Central Feature	No
o Directional Distribution	Yes
o Linear Directional Mean	No
o Mean Center	Yes
o Standard Distance	Yes
❊ Utilities	
o Calculate Area	No
o Collect Events	Yes
o Count Rendering	No
o Export Feature Attribute to ASCII	No
o Z-score Rendering	Yes
4. Geocoding Tools	No
5. Data Interoperability Tools	No
6. Data Management Tools	
❊ Database	No
❊ Domain	No
❊ Feature Class	
o Calculate Default Cluster Tolerance	Yes
o Calculate Default Spatial Grid Index	Yes
o Integrate	Yes
❊ Feature	
o Add XY coordinate	No
o Check Geometry	Yes
❊ General	
o Merge	Yes
o Copy	Yes
o Append	No
❊ Fields	
o Add Field	Yes
o Calculate Field	Yes
❊ Generalization	
o Dissolve	No
❊ Indexes	
o Add Attribute Index	Yes
❊ Join	
o Add Join	Yes
❊ Layers and Tables View	
o Make Feature Layer	Yes
o Make Query Table	Yes
o Make XY Event Layer	Yes
o Select Layer by Location	Yes
o Select Layer by Attribute	Yes
7. Cartography Tools	
❊ Masking Tools	
o Cul-De-Sac Masks	No
o Feature Outline Masks	Yes
o Intersecting Layer Masks	Yes
8. Network Analyst Tools	No
9. 3-D Analyst Tool	No
10. Geostatistical Analyst Tools	No
11. Spatial Analyst Tools	No

About the Authors

Kevin M. Curtin is Associate Professor of Geography at George Mason University. His teaching and research interests include Network GIS, Transportation Geography, and Optimal Facilities Location Science.

Corresponding Address:
Department of Geography
George Mason University
4400 University Drive MS1E2
Fairfax, VA 22030
curtin@gmu.edu

Greta Nicoara has a master's degree in Geospatial Information Sciences from the University of Texas at Dallas. She is currently a Senior GIS Analyst/Programmer with the City of Richardson, Texas.

Corresponding Address:
Rumana Reaz Arifin
Program in Geographic Information Sciences
School of Economic, Political and Policy Sciences (GR 31)
University of Texas at Dallas
800 West Campbell Road
Richardson, TX 75080-3021

References

Curtin, K. M. 2007. Network data structures. Encyclopedia of geographic information science. K. Kemp, ed. Thousand Oaks: Sage Publications.

Curtin, K. M. 2007. Network modeling. The encyclopedia of geoinformatics. H.

Karimi, ed. Hershey: Idea Group Publishing.

Curtin, K. M., V. Noronha, et al. 2001. ArcGIS transportation data model. Redlands, CA: Environmental Systems Research Institute.

Dueker, K. J., and J. A. Butler. 2000. A geographic information system framework for transportation data sharing. Transportation Research Part C-Emerging Technologies 8(1-6): 13-36.

Dueker, K. J., and R. Vrana. 1992. Dynamic segmentation revisited: a milepoint linear data model. Journal of the Urban and Regional Information Systems Association 4(2): 94-105.

ESRI. 2001. Linear referencing and dynamic segmentation in ArcGIS 8.1. Redlands, CA: 56

ESRI. 2003. Linear referencing in ArcGIS: practical considerations for the development of an enterprise-wide GIS. Redlands, CA.

Federal Highway Administration. 2001. Implementation of GIS based highway safety analysis: bridging the gap. McLean, VA: U.S. Department of Transportation, 104.

Federal Transit Administration. 2003. Best practices for using geographic data in transit: a location referencing guidebook. Washington, D.C.: U.S. Department of Transportation, 211.

Fletcher, D., J. Espinoza, et al. 1998. The case for a unified linear reference system. *Journal of the Urban and Regional Information Systems Association* 10(1).

Fohl, P., K. M. Curtin, et al. 1996. A Non-planar, Lane-based, Navigable Data Model for Intelligent Transportation Systems. International Symposium on Spatial Data Handling, Delft, The Netherlands, International Geographical Union.

Goodman, J. E. 2001, November 28, 2001. Maps in the fast lane—linear referencing and dynamic segmentation. April 23, 2004, http://www.directionsmag.com/article.php?article_id=126.

Kiel, D., J. Pollack, et al. 1998. Issues in adapting linear referencing systems for transportation applications: current practice and future outlook. Journal of Computing in Civil Engineering 12(2): 60-61.

Koncz, N. A., and T. M. Adams. 2002. A data model for multi-dimensional transportation applications. International Journal of Geographical Information Science 16(6): 551-69.

Koncz, N. A., and T. M. Adams. 2002. A data model for multi-dimensional transportation location referencing systems. Journal of the Urban and Regional Information Systems Association 14(2): 27-41.

Koncz, N. A., and T. M. Adams. 2002. Temporal data constructs for multidimensional transportation geographic information system applications. Transportation Research Record 1804: 196-204.

Miller, H. J., and S. Shaw. 2001. Geographic information systems for transportation. New York: Oxford University Press.

Noronha, V., and R. L. Church. 2002. Linear referencing and other forms of location expression for transportation. Santa Barbara: Vehicle Intelligence & Transportation Analysis Laboratory, University of California, 26.

Nyerges, T. L. 1990. Locational referencing and highway segmentation in a geographic information system. ITE Journal, March: 27-31.

Scarponcini, P. 2001. Linear reference system for life-cycle integration. Journal of Computing in Civil Engineering 15(1): 81-88.

Scarponcini, P. 2002. Generalized model for linear referencing in transportation. Geoinformatica 6(1): 35-55.

Sutton, J. C., and M. M. Wyman. 2000. Dynamic location: an iconic model to synchronize temporal and spatial transportation data. Transportation Research Part C-Emerging Technologies 8(1-6): 37-52.

Texas Department of Transportation. 2003. Texas reference marker (TRM) system user's manual, 493.

Vonderohe, A., T. Adams, et al. 1998. Development of system and application architectures for geographic information systems in transportation. Washington, D.C.: National Cooperative Highway Research Program, Transportation Research Board, 23.

Vonderohe, A., C. Chou, et al. 1997. A generic data model for linear referencing systems. Washington D.C.: National Cooperative Highway Research Program, Transportation Research Board.