342

Map accuracy and location expression in transportation – reality and prospects

Val Noronha *, Michael F. Goodchild [1]

Vehicle Intelligence and Transportation Analysis Laboratory, NCGIA, University of California, Santa Barbara, CA 93106-4060, USA

Abstract

We were contracted to test a suite of proposed location messaging standards for the intelligent transportation systems (ITS) industry. We studied six different databases for the County of Santa Barbara, documented types and magnitudes of error, and examined the likely success of the proposed standards. This paper synthesizes the test results and identifies caveats for the user community as well as challenges to academia. We conclude that, first, current messaging proposals are inadequate, and superior methods are required to convey both location and a measure of confidence to the recipient. Second, there is a need to develop methods to correct map data geometrically, so that location is more accurately captured, stored and communicated, particularly in mission critical applications such as emergency servicing. To address this, we have developed methods for comparing maps and adjusting them in real time. Third, there must be standards for centerline map accuracy, that reflect the data models and functions associated with transportation. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Location referencing; Intelligent transportation systems; Geographic information systems; Street network databases; Map database interoperability; Transportation datums

1. Introduction

Although it is several years, even decades since the advent of digital maps, global positioning systems (GPS) technology and machine-searchable street names and coordinates, it can be surprisingly difficult for the average person to describe a location, even one limited to the discrete

* Corresponding author. Tel.: +1-805-893-8992.
  E-mail addresses: noronha@ncgia.ucsb.edu (V. Noronha), good@ncgia.ucsb.edu (M.F. Goodchild).
  [1] Tel.: +1-805-893-8049.

confines of a street network. This problem is faced daily by people reporting accidents and vehicle breakdowns, and by the emergency service personnel receiving and servicing those calls.

Recent technological development in intelligent transportation systems (ITS) allows vehicles and pedestrians to be tracked in real time, using GPS, cellular phone tower triangulation, inductive loop detectors embedded in highway pavement, or closed circuit television (CCTV) cameras; even satellite imagery is being proposed as a tracking option. The methods have applications in emergency servicing, real-time highway information provision (e.g. congestion, construction, fog) and traffic management, hazardous material management, travel demand studies, law enforcement and criminology. They are certain to generate a large volume of geographically referenced data in the coming years. But given the inherent difficulty in describing location, the positional references in such data could be ambiguous or erroneous, and the errors could propagate through subsequent processing of the data.

In the United States, ITS is a major area of technology research, its development abetted by two transportation bills, the 1991 Intermodal Surface Transportation Efficiency Act (ISTEA), and the 1998 Transportation Equity Act for the 21st century (TEA-21). ISTEA spurred the initial development of ITS concepts, notably the national ITS architecture (USDOT, 1997), but it became clear by the mid-1990s that significant operational problems remained to be resolved. One was the need for standards for interoperability, particularly in location referencing. TEA-21 explicitly addressed this, and in 1998 the US Department of Transportation released a list of "critical standards" (USDOT, 1998) for immediate research, development and testing. Interoperability in general, and location referencing in particular, featured prominently in this list.

Emergency management services (EMS) is an important component of ITS, because they present challenges similar to incident management (IM) services in ITS, and because the road system is the medium of delivery of service to most non-road emergencies. The problem of determining the precise location of an incident, identification of appropriate resources, facilitation of service delivery by signal pre-emption, and real-time traffic management or evacuation, are issues that cut across the boundaries of EMS and ITS.

Another aspect of transportation research, with potentially large impacts on infrastructure, is the periodic systematic study of travel behavior, best exemplified by the Federal Highway Administration's Nationwide Personal Transportation Survey (NPTS). Individuals are asked to maintain diaries over a period of days or weeks, recording each trip they make for personal or business reasons, listing time of day, origin and destination, intervening stops and route. Battelle (1997) conducted an experimental variant on the standard methodology in Lexington, Kentucky, using GPS loggers to supplement the respondents' answers; similar methods have since been explored in Georgia and Texas. A study by the California Air Resources Board tracked trucks with GPS to study their traffic patterns, and their implications in air quality. Such new technology solutions are obviously data rich, but there is the potential for locations – origins, destinations and routes – referenced in the data to be incorrectly interpreted due to GPS and map data errors. Clearly the need for accuracy is greater in some applications than in others.

Problems with location reporting have been observed anecdotally for a number of years, but there has been relatively little scientific documentation of the scale of the problem and its solution. Some researchers have focused on the problem of address matching and geocoding (e.g. Kim and Nitz, 1994), and a number of studies have examined GPS error independently (e.g. Quiroga and

Bullock, 1998). Neither of these addresses the broader issues of location reporting in a transportation context.

This paper examines these issues analytically, synthesizing results of formal tests conducted by the Vehicle Intelligence and Transportation Analysis Laboratory (VITAL) in response to the need for standardization in the ITS industry. It expands on ideas by Noronha (1999). Section 2 presents the need for location expression and exchange with particular reference to ITS scenarios. Section 3 examines the types of error commonly found in contemporary digital maps, and their impact on location referencing. Section 4 discusses short- and long-term measures to improve the accuracy of location reporting. The paper concludes with speculations on future requirements, and the academic, technical and managerial challenges that lie ahead.

2. Location expression and the interoperability problem

For transportation purposes, location referencing is the description of location of a static or dynamic object, in one, two or three dimensions. In the case of dynamic objects there may be reference to a fourth dimension. The object may be a disabled vehicle, which is static and has a relatively easily defined location. Alternately, it may be an incidence of congestion, the boundaries of which are not easily defined, though if modeled appropriately, congestion can be sampled using loop detectors and CCTV. A toxic plume is an example of an object that is dynamic, that may have poorly defined boundaries, and whose location is not easily measured. This paper does not attempt to deal with these difficult cases; it addresses the relatively simple instances of disabled vehicles and similar point incidents.

ITS scenarios require cooperative action between parties, each operating in a free market with respect to proprietary databases and systems. Take an example of a vehicle collision that results in casualties, on an urban freeway interchange during rush hour. Numerous agencies, both public and private, are involved in the response – police, fire engines, ambulances and tow trucks. Traffic management centers (TMCs) and private information service providers (ISPs) monitor the incident, advise motorists and divert traffic. Each agency employs an internal communications code that is meaningful in the context of its operations. GPS users quote coordinates, tow truck operators use intersecting street names, highway maintenance crews operate with respect to linear measures from a defined starting point, some may apply an undocumented nickname such as "Suicide Bend" to an accident-prone section. No matter what the language, agencies must understand the precise location of the incident, its location with respect to other ramps and roads, and the dynamics of its impact on traffic conditions. The purpose of standardization is to ensure that parties communicate unambiguously and error-free with respect to mission-critical information items.

We use the term location expression (LX) to mean the description of location by coordinates, street names or other means. The process of communicating an LX is location expression exchange, or LXX. An object that changes location over time, such as a moving vehicle, is described by multiple pairings of an LX with a time expression, (TX) (note that time is also subject to uncertainty and variation in expression, e.g. "1430 h PST", "yesterday" or "every Tuesday in the summer"). The LX/LXX nomenclature isolates the components of the location reporting problem. Our concern with LXX is specifically in the context of machine-to-machine communication,

where human intervention and interpretation are not possible. ITS scenarios of vehicles making emergency calls on behalf of their unconscious drivers rely on such exclusively machine-based interaction.

At first sight, the transformation between LXs, from coordinates to street names, linear or grid references, is a simple geometric process; even if agencies work with different coordinate systems, such as Universal Transverse Mercator (UTM) and State Plane projections, translations are well defined mathematically and relatively easily achieved. In reality, there are two problems that conspire to make LXX an error-prone process. The first is the accuracy of the initial position capture and expression. The highway incident above is typically reported to emergency services by several cellular phone callers, their descriptions perhaps pointing to different locations. Police often record accident data at the nearest intersection to an incident, because they do not have the means to measure the distance from the intersection. Even if the incident location were reported by GPS coordinates, the 10–100 m GPS error could position the incident on the wrong highway ramp or on an adjacent service road. The second problem is the error in the reference map with which the location is interpreted. This is a complex problem, and is discussed in detail in the next section.

3. Error in map databases

To understand the extent of the spatial interoperability problem, one must appreciate the types and degree of error in digital map databases. In general there are errors in position, classification and inclusion, names and other descriptive attributes, linear measurement and topological relationships. The following discussion is based on tests conducted in the County of Santa Barbara, California, using six databases from public and private sources, representing the principal street map vendors for ITS in the US. Although we use the term "error" liberally in this discussion, bear in mind that these databases were developed for different purposes, from different scales of survey and mapping, and that their sale prices vary from $1500 to $45 000. We do not identify vendors or compare products, and our comments are intended not to be critical of vendors, but to illustrate the kinds of operational hurdles that exist and must be overcome by users.

Testing involves sampling a set of point locations, either lab-generated or selected in the field (field locations are associated with landmarks such as utility poles, and documented by notes and photographs). In some cases, sample points are generated at regular or random intervals along network polylines; in other tests, they are required to be at commonly identifiable locations such as intersections. Major streets are examined separately, on the grounds that initial ITS deployment would be on highways and major arterials only. The sample points are transferred to each database in turn by the LX method under test, and an appropriate measure of success applied to examine the accuracy of the transfer. The test results are obviously specific to the County of Santa Barbara and the time of acquisition of the data sets (May, 1997). The county does present a representative and average cross-section of street styles, both urban freeways and sinuous mountain roads, but it is reasonable to believe that results would be better with databases focused on large urban centers, and conversely poorer in more remote locations.

3.1. Position

Although digital maps are sometimes thought to be exceptionally precise, there is no reason why they should be better than hardcopy products. Whether paper or digital, a map is an iconic representation of a reality that can be interpreted in different ways. That reality itself may not be constant. For example, street centerlines and intersections are defined precisely at the time of construction planning, but once the survey stakes are removed and the asphalt laid, the centerline is represented by a stripe of paint, which wanders slightly each time there is construction activity or re-striping. Consequently there is about ±0.5 m of lateral uncertainty in a centerline, and ±2–3 m in the definition of a large intersection. A gore point (the "V" at a highway exit, where the right shoulder stripe of the highway and the left shoulder stripe of the exit meet) may wander about 10 m due to striping imprecision – this is not immediately evident from the vantage point of a moving vehicle. These are examples of uncertainty in a real-world object; in addition there is error in the measurement and representation of that object. The earth's frame of global reference is itself subject to some uncertainty due to wobbling, tectonic movement and tides, and true sub-metre accuracy is achievable only in relation to local monuments. Digitization of a centerline from aerial photographs or older maps inevitably introduces error and generalization, the amount depending on the source scale. Some vendors evolved into the digital era after long-established practices in the hardcopy business. For the sake of clarity or esthetics, they had deliberately distorted the alignment of some roads such as service roads running parallel to highways. Or they had taken the position – valid in the pre-GPS era – that absolute positional accuracy was unimportant as long as the map represented names, gross shapes and topological relationships accurately. The cumulative effect of these sources of uncertainty and error is that road alignments can differ as much as 100 m in urban areas, and 200 m on winding mountain roads. An incident at a freeway interchange can therefore appear to be on the wrong ramp (Fig. 1(a)), or in an adjacent neighborhood (Fig. 1(b)). It is unlikely that a third coordinate (elevation) can resolve this reliably,
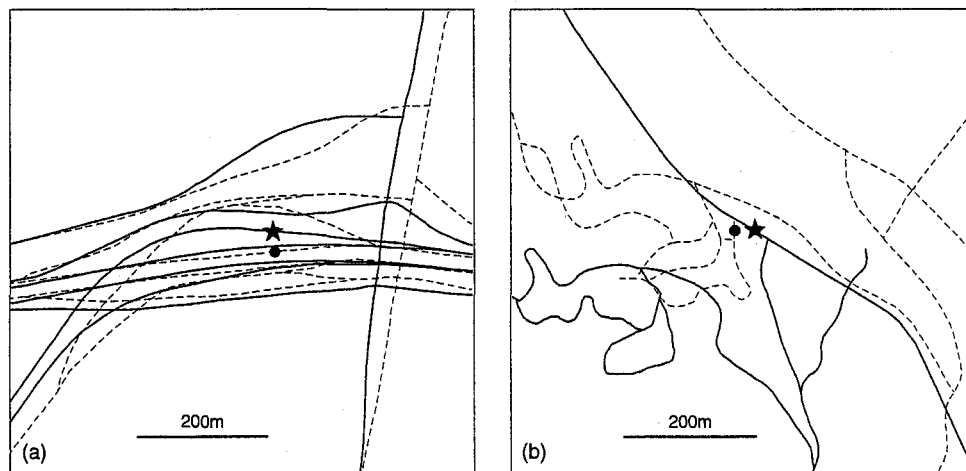

Fig. 1. Positional error. A coordinate from a vehicle on a ramp or highway in map A (star on solid line) snaps to the wrong ramp or street in map B (dot on broken line): (a) shows a freeway interchange in Santa Barbara; (b) is a highway and adjacent neighborhood on the outskirts of the city.

for two reasons: first, current databases do not contain the third dimension; secondly, error GPS elevation readings is generally twice as high as horizontal error.

To correct error, one must first measure, understand and model it. For a single point, obvious measure of error is the Euclidean distance between the point on the erroneous map, a₁
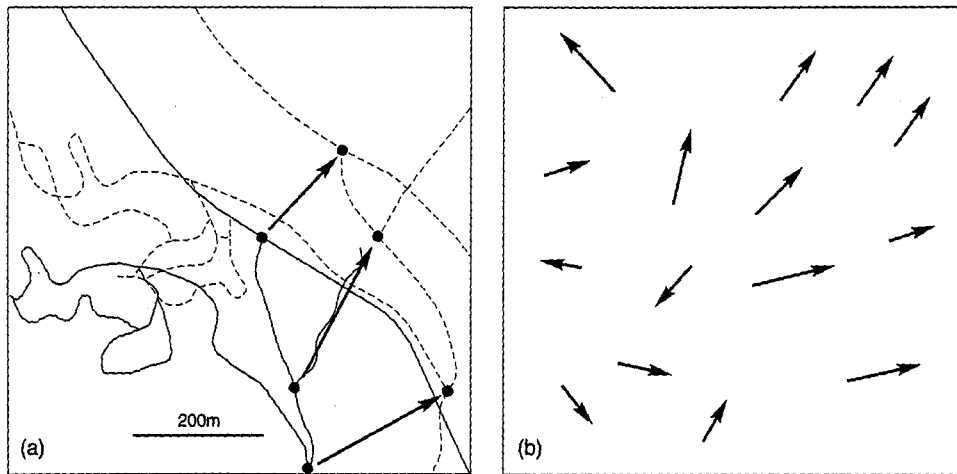

Fig. 2. (a) Error vectors constructed between points on map A and corresponding points on map B. (b) Vector field. Errors that are consistent in magnitude and direction (northeast of figure) are easy to model and to correct. When vectors are inconsistent (western portion of figure), extensive correction or complete re-survey may be required.
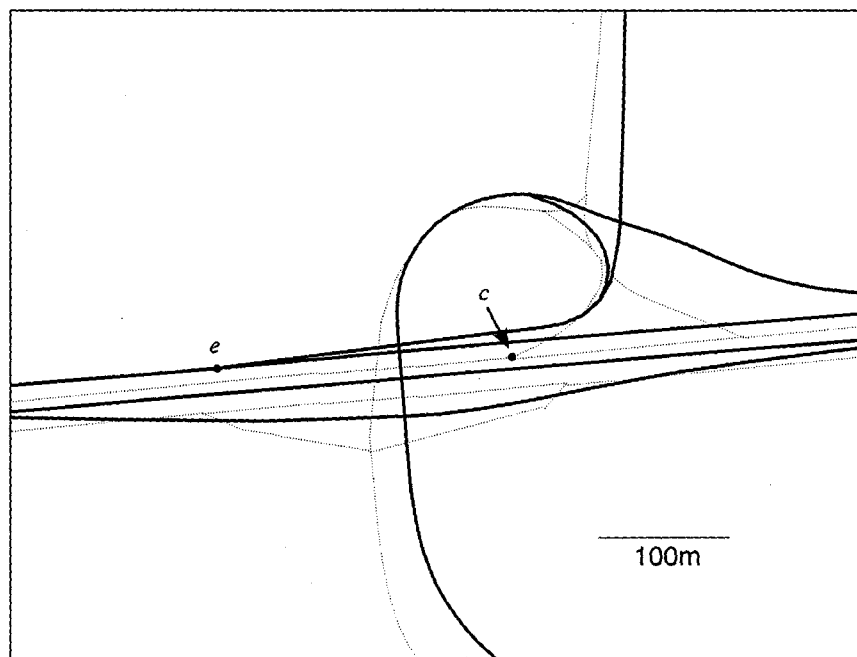

Fig. 3. Overlay of two maps showing different positions (c and e) where ramp meets freeway. When these are represented by planar-enforced models, topological relations clearly differ. The impact on linear measurements taken from this point is more than 200 m. Reproduced from Noronha (1999) with permission.

the true position of that point. The vectors between the points (Fig. 2(a)) show the magnitude as well as the direction of the error. Note that the term error assumes that one version is accurate and the other is not; this may not be easy to determine, considering the remarks in the last paragraph. Our analysis of error is based on an engineering scale map of the study area, and verification by differential GPS. Matching of points from one database to another is accomplished by a combination of automated and manual methods.

Extending the vector analysis to a larger area, one can develop a vector field (Goodchild and Hunter, 1997) that shows the variation in error over space (Fig. 2(b)), and suggests its origins. One could hypothesize that the field is continuous, in which case the error at a given point can be interpolated from that at surrounding points; our preliminary studies indicate that this is true at least to some degree (Church et al., 1998; Funk et al., 1998). Error direction and magnitude are often consistent over entire neighborhoods, but they change abruptly across some boundaries. This indicates that the map was developed piecemeal, at different times, or from different sources. Continuity of the error field also breaks down on freeways. Intersections of ramps with freeways are prone to high positional error that is difficult to model, because the paths intersect at a small angle, and a slight error in alignment translates into a large longitudinal displacement (Fig. 3).

3.2. Scale, classification, inclusion and topology

It was argued above that mapping necessarily entails an interpretation of reality. This is truest in the case of road inclusion and classification, which are closely connected with map scale and resolution. Dual carriageway freeways are single lines at one scale and double lines at larger map scales. Similarly there are differences in the treatment of traffic circles, median strips and channelized turn lanes. It is impossible to find two maps that agree on their categorization of "major" roads, let alone the fine distinctions between arterials and collectors. Discrepancies in classification lead to differences in inclusion. Driveways into condominiums and building complexes (e.g. hospitals) do not appear in all maps, because not all vendors consider these part of the street network. Inclusion is also impacted by the effectiveness of map update. The vendor that promptly includes new neighborhoods clearly shows several streets that another vendor does not.

A further consequence of discrepancies in inclusion is conflicts in topology (the term is used in the GIS sense, of connective relationships between points, lines and polygons). Clearly if one map shows a number of driveways off a principal road, while another map ignores them, the topological relations differ. Another aspect of topological inconsistency is the reduction of non-planar intersections (i.e. overpasses). GIS structures modeled on the TIGER database (Marx, 1986) impose planarity on all intersections due to the need to consider roads as potential polygon boundaries. Navigation turn tables are built into the database structure, with a large artificial impedance on turns that are not physically possible. Problems then arise when gross topological errors occur in the database, e.g. when a cloverleaf ramp meets a freeway east rather than west of an overpass (Fig. 3). Correction of the positional error requires extensive recoding of the topological relationships within the interchange.

It is tempting to study some of the above differences in inclusion and classification by characterizing each road section by some combination of name and position, and looking for a corresponding segment in the other database. This does not work because of (a) differences in naming, discussed in detail below, and (b) topological discrepancies that result in one map having

long sections of road, while another fragments them into short segments. Nystuen et al. (1997) study both positional and inclusion differences between maps by buffering the centerlines in each map, and topologically intersecting the buffers. The degree of correspondence is analyzed spatially by gridding the area and measuring the intersections between buffers, cell by cell. This method is attractive because it mimics visual analysis and is intuitively valid; a drawback is that finer topological differences such as dual-line versus single-line highways do not affect the statistics appreciably, although these differences can have significant impacts on interoperability.

3.3. Names

In the light of the inadequacy of coordinates as an unambiguous LX, it has been suggested that street names, in the form of street addresses or intersections, would be more reliable. This is a reasonable argument up to a point – after all, the post office delivers mail based on street addresses where they exist, with little error. There are two caveats to this. First, mail delivery is facilitated by postal codes that enable sorting and reduce ambiguity, and also by intelligent human interpretation that easily substitutes an incorrect "Market Street" with the correct "Marquette Street". Secondly, transportation applications are not restricted to mail delivery routes. Many roads (notably freeway ramps) do not have names; for Santa Barbara County as a whole, 20–45% of records have blank name fields.

Further problems exist in the capture and storage of names. Abbreviations are inconsistent (e.g. "Av" versus "Ave"), parsing of fields varies (e.g. Spanish street types such as "Via del" should be separated as street type prefixes, but they are often incorporated into the proper name), and only one of our six databases provides for aliases that would equate "San Diego Freeway" with Interstate 5. There are consistency errors even within databases: one section of a road is correctly named "Winding Way" while another contains a typographic error, "Winting Way". Appropriate data structuring would avoid this.

Table 1
Four stretches of road, as named in five databases A–E[a]

Road	Database				
	A	B	C	D	E
1	E Camino Cielo	Camino Cielo	[blank]	[Street not present]	E Cm Cielo
2	San Marcos Pass Hwy 154	Highway 154	[blank]	San Marcos Pass CA-154	San Marcos Pass [blank]
3	Mountain E Mountain	Mountain Park Bella Vista	[blank]	Mountain	W Mountain E Mountain Park Bella Vista
4	Foothill Cathedral Oaks	Cathedral Oaks	Hwy 192	Foothill Cathedral Oaks	Foothill Cathedral Oaks

[a] A road may have several names over the sampled distance; databases may not agree on where changes occur. Only A provides for an alias. Reproduced from Noronha (1999) with permission.

Table 1 shows four randomly selected lengths of street, as named in five databases. The streets are selected visually, and some are so long as to have multiple names over their course. The variation in names offers a quick insight into the matching problem.

For the reasons outlined above, success rates in matching names across databases are lower in ITS applications than even the 60–75% often reported for address matching as used with street addresses. Intelligent software does exist to deal with some of the problems, e.g. to equate Winding with Winting, Market with Marquette; and even CA-101 with HWY-101 (Fellegi and Sunter, 1969; Knuth, 1973; Jaro, 1989); but software alone cannot resolve the more difficult cases, particularly where blank names are encountered.

3.4. Linear measurement

Prior to the advent of GPS, the only practical way to describe a location on a remote section of road was by linear measurement from a defined starting point. Transportation engineers continue to use linear references today, sometimes to the exclusion of maps and two-dimensional referencing. While motorists' experience of linear referencing is limited to the 100 m resolution of vehicle odometers, professionals employ distance measuring instruments (DMI) for 1 m-resolution readings. DMI technology has evolved over the years, from mechanical or optical revolution counters attached to the wheels of a vehicle, to modern electronic pulse sensors that hook into the transmission.

There are three components to a linear reference: (a) the road on which the point lies, described by a road name or numeric identifier; (b) the origin and direction for the measurement; (c) the distance measurement. Error in (a) is eliminated if parties communicate using standard identifiers, else the comments on street name interoperability above apply. Specification of the origin (b) is similarly susceptible to identification error. The origin has to be described with precision equivalent to that of the measurement: if the distance is stated with 1 m precision, it is not sufficient to specify the origin simply as the intersection of two streets, when the intersection is 30 m wide. Finally the measurement (c) is subject to errors, depending on curvature and other physical characteristics of the road. Throughout this paper we assume offsets to be measured along the road (in the case of the DMI) or along its polyline representation.

The linear LXX question is two-fold: (i) how well a location expressed by other means (e.g. coordinates or street names) translates to a linear reference, and (ii) how well a linear reference measured by a DMI, or with respect to one map database, describes that location with reference to a different map database. Regarding (i), in the Santa Barbara databases, there is on an average a 40% chance that a GPS point (with 100 m selective availability error) snaps to the wrong road, and a further 10% chance that it snaps to the wrong topological section of the correct road. The average linear reference error from such a point is about 90 m, assuming no recalibration points (VITAL, 1999). If the point is specified directly from DMI measurement, there are errors due to instrument limitations, calibration and operation, usually the greater of 1% or 2–10 m. On the matter of (ii), the relationship between polyline geometry, generalization and linear measurement has been studied from a number of standpoints (Mandelbrot, 1967; Douglas and Poiker, 1973; Buttenfield, 1985). It is generally well known that as a polyline is subjected to increasing generalization, the digitized length decreases. The Santa Barbara data set shows average differences of 8% in digitized length (longest versus shortest version of the same road) on a variety of roads, with

a range of 1–16%. The deviation of digitized length from DMI length for a given data set over all sample roads ranges from 0.5% to 4% with an average of 3% and a worst case of 16%. Compensation for this error can be applied by normalizing a linear offset, i.e. expressing the offset as the proportion of its distance from the origin, to the total length of the road section (clearly this adds to the cost of specifying the location, because it requires two measurements to be made). Absolute linear transfer errors are about 50 m on average, and normalized transfer errors are about 25 m; errors up to nearly 1 km are encountered in exceptional cases, due to longitudinal errors in defining the intersections of ramps with highways (VITAL, 1999).

4. Error remediation

Given the immediate need for incident reporting and amelioration of transportation problems, it is necessary to examine strategies that enable the deployment of services using current databases. Taking an ITS example, a motorist may buy an in-vehicle database from vendor A, receive data from a TMC that operates with vendor B, and pass messages to an emergency center that uses vendor C. Regardless of which pair of databases is involved in a transaction, the message must be received unambiguously and error-free.

In addition to ITS concerns, there is the general problem of backward compatibility of databases. A wealth of legacy data is attached to databases of differing positional quality. For example, an emergency center may attach a note to a 1980s-vintage database, that 872 Whistler Highway is the farm of John Brown, with a large red barn near the driveway. It is important that when the emergency center upgrades to a more current database, the details of the Brown farm be preserved. This section examines how this might be done.

4.1. Interoperability standards

Fletcher (1999) defines four levels of interaction, the "Four I's": (a) integration, in which parties share hardware and software from a common vendor, or use a single database in which internal identifiers are universally understood, e.g. within a small office; (b) interoperability, where vendors may differ, but the semantics of objects and processes are standardized, and systems differ only in the internal details of implementation; (c) interfacing, where there are differences in semantics, traditions, systems and vendors, and communication is achieved by means of third-party translators that attempt to harmonize semantics to the extent possible; and (d) independence, in which parties fail to communicate because their systems are radically different. Interfacing is generally considered less elegant than is interoperation, and attracts irreverent terms such as "duct tape". But given the considerable problems with map databases, we may not have the luxury of restricting ourselves to any one of these I's, hence the word interoperability is used in the broadest sense below.

4.1.1. Messaging

One obvious candidate for standardization is the message used to communicate a location. In the language of the Open Systems Interconnect (OSI) model (Day and Zimmerman, 1983), this discussion of message standardization pertains to the upper level layers, application and

presentation. In the United States the Society of Automotive Engineers (SAE) publishes J2374, the national Location Referencing Message Specification (SAE, 1998), which currently has the status of an "information report" rather than a standard. J2374 is a proposed set of seven message profiles to communicate a location, each profile generally being a form of LX. A user group chooses a profile to suit its mode of operations, for example a location derived from GPS is usually transmitted using a coordinate profile, whereas a linear GIS-T application employs a linear referencing profile. J2374 does not specify how the message is to be interpreted at the receiving end; it does not contain metadata or any explicit measure of message reliability; and it provides only for one-way communication – there is no provision for the originator to know whether the message has been successfully received and correctly interpreted.

The cross-streets profile (XSP) is the J2374 profile proposed for widest use in communicating between TMCs, and from TMCs or ISPs to vehicles; it is also the basis of SAE's J1746, the ISP-to-vehicle standard. The original XSP described a location in terms of: (a) the name of the principal street on which the location lies; (b) the names of the two cross-streets that bound the segment of the principal street; and (c) the linear offset along the principal street, measured from one of the cross-streets (Fig. 4). At first sight this appears to be an adequate way to represent location, but on close examination, and bearing in mind the context of machine-to-machine communications, there are several problems. First, there is no provision in the XSP for geographic region, so a matching triad of street names anywhere in the world is a potential candidate. Second, due to the high incidence of blank name fields in current databases, a XSP message can be composed successfully (i.e. with all three street names non-blank) in only 33% of test cases. Third, when certain street configurations are encountered (e.g. crescents, intertwining roads with multiple crossings, closed loops) the logic of the XSP breaks down. In sum, due to these and other problems with name matching outlined above, the original XSP is successful (i.e. accurate and unambiguous) only in about 25% of all cases – this applies even to major roads, because highway ramps are unnamed. Following our original suite of tests (VITAL, 1999) we recommended that the XSP be strengthened by including a pair of coordinates in the profile, to establish general location. The measure of success is now complex, because the enhanced XSP contains two entirely independent ways of expressing location, which could be in conflict; or there could be several partial but different fuzzy lexical matches with the cross-street triad, which would have to be ranked by some rule based on the strength of lexical match and distance from the coordinates. In general, the success rate for the enhanced XSP is 50–80%. This still leaves a large proportion of cases that fail, or are ambiguous. Details on the tests are reported by VITAL (1999) and Noronha et al. (1999).
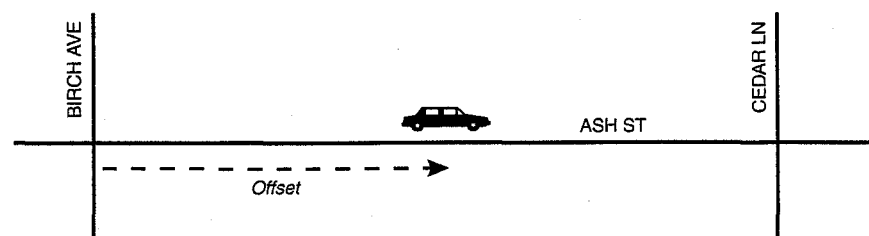

Fig. 4. The J2374 cross-street profile expresses location in terms of a principal street (Ash), two cross-streets (Birch and Cedar) and the offset distance from the first cross-street.

An obvious need in messaging is for the user to have a measure of information reliability, and it may be important for the sender to know whether the message has been satisfactorily received. VITAL has therefore proposed a robust, theoretically 100% successful protocol termed "LX-100", for use in mission-critical LXX. LX-100 mimics the negotiation process that takes place in verbal communication when one describes an emergency or any other location. Given the luxury of time, say to describe the venue of an office party, one employs a combination of street names, landmarks, absolute and relative navigation directives, which are together redundant; the receiver processes the information, detects success or failure based on redundant clues, and requests additional information in the event of failure. Similar principles are used for error checking in digital packet data transfer, around which the Internet Protocol is built; the European standards development and testing initiative, Extensive Validation of Identification Concepts in Europe (EVIDENCE), also uses redundant LX in its concept of the intersection location (ILOC).

LX-100 is still under development, and many issues and algorithms remain to be resolved. If the process is to approach 100% success, it must make assumptions about the quality and fitness for use of reference databases.

4.1.2. Database quality

The issue of database quality standards is sensitive, because it directly impacts the operations and financial viability of data vendors – and standards can be successful only if players subscribe to them. It takes large investments to build high-quality network databases, and if the benefits are expected to accrue only at some undefined time in the future, data vendors understandably channel resources only into areas of unquestionable benefit. Therefore a plea for standards must be cast in an incremental and economically feasible framework.

One obvious and realizable area for standardization is street naming, especially abbreviation, parsing of fields, and treatment of shared names (aliases such as San Diego Freeway and I-5, as well as cases where highways bear more than one designation when they merge temporarily). It is essential that standards govern the naming of highways and ramps, particularly because ramps form the topological intersections that define highway segments. There is also a need, though not as urgent, to employ classification standards based on road width and traffic volume, so that arterial roads are defined consistently across databases. All these are relatively easy to achieve, and would greatly improve interoperability in the short term.

The more challenging issues for standardization are those related to map scale, resolution, inclusion, topology, currency, coordinate and attribute accuracy. Compliance with some of these would require extensive field work and re-survey, and perhaps entirely new data models. Resolution standards have to specify that divided highways, traffic circles and channelized lanes should be explicitly represented if they meet certain criteria of size, separation and accessibility. Topological and data structuring standards must ensure that non-planar intersections are appropriately represented. Coordinate accuracy is a function of map scale, and a vendor could argue that a map meets standards for 1:100 000 mapping, although the corresponding ±50 m accuracy may not be sufficient for some applications. There may be a need to certify a product's fitness for use in an application.

4.1.3. Datums

The more ambitious aspects of standardization, such as coordinate accuracy, may not be realizable by private vendors alone, and government invention may be required. In the 1980s and

1990s, the ready availability of DIME and TIGER files in the US spawned a large number of value-added products; similarly it may be that a high-level effort to standardize coordinates may be an appropriate government initiative in the 21st Century. This does not necessarily require re-survey and development of a new high-quality data base: a skeleton of well defined points will suffice in most areas. The ITS Datum (Siegel et al., 1996) is a proposed set of monuments designed to serve map accuracy needs in ITS, in both the short and long terms. In the short term, it provides an interim vehicle for vendors to improve the accuracy and interoperability of their products; in the long term it offers a unifying set of reference points for data exchange, the benefits of which need not be limited to ITS.

Preliminary study at VITAL indicates that the accuracy of ITS Datum points would vary with the needs of applications (Table 2), but would generally be defined to ±3–5m. Density would be higher in areas of greater road density, for example, each ramp in a freeway interchange would require at least one Datum point, probably at the intersection of the ramp with the freeway. A test database of about 5000 points has been compiled, and a rubber-streeting™ algorithm developed to match one data set to another geometrically, in real time (Fig. 5). The algorithm works satis-factorily on the test area shown, but it is now clear that mere specification of some points with accurate coordinates, identifiers and segment lengths will not suffice, and further work remains to be done on the development of appropriate data models. If the Datum is properly designed and implemented, there will be no need for manual identification of points of correspondence between databases, as was required to derive the experimental results reported above.

Two other national datums are currently being discussed: a linear datum (Vonderohe et al., 1995) to improve accuracy of linear measurement in GIS-T, and a road identification standard (FGDC, 1998), to enable interoperation of public sector databases as part of the National Spatial Database Infrastructure (NSDI). The work on the linear datum is described elsewhere in this book. It proposes a conceptual network of anchor sections, segments of road with precisely de-fined lengths. Anchor sections terminate at anchor points, which are defined by recoverable field locations (not necessarily intersections) that are selected solely to achieve the objective of opti-mizing linear accuracy. The NSDI proposes a set of points and line segments that are readily identifiable in the field and in databases (e.g. intersections), to which standard identifiers are applied, that serves as public keys to enable data exchange. The design enables data sharing between dual-line and single-line representations of the same road, and largely considers two-or three-dimensional positional accuracy to be irrelevant because interoperability is achieved by the use of standard identifiers, not positional similarity between objects. Clearly there are parallels

Table 2
Error tolerance in ITS applications, present and future[a]

Period	Locale	Lateral tolerance (m)	Longitudinal tolerance (m)
Present (2000)	Open highway	25	50–100
	Urban/interchange	5	50
Future (2015)	Open highway	1	10–20
	Urban/interchange	1	5

[a] The numbers are coarse estimates; the absolute values are less important than the relative values and the structure of the table.

Fig. 5. VITAL's rubber-streeting algorithm applies a geometric adjustment to one map to make it agree with another more accurately surveyed map or datum. Reproduced from Noronha (1999) with permission.

between these concepts and the ITS Datum, which incorporates all the requirements of the other standards – linear accuracy, standard identifiers – and three-dimensional accuracy.

4.2. Other solutions to the interoperability challenge

VITAL (1999) has shown that while some of the interoperability test results with current da-tabases are poor, results with differential GPS and better quality databases are excellent, at least

for current application requirements. One might legitimately argue that new surveys will soon brighten ITS interoperability prospects. While this is true of the future, there are two causes for caution. First, GPS-based re-survey is technologically possible today, but remains largely undone because of cost. An alternate way to improve public sector databases is by integration of construction and engineering data into transportation GIS; but for technical and management reasons, there are few jurisdictions where this has been done effectively. Second, improved coordinate accuracy is only one aspect of quality. Legacy attribute data have to be merged with new coordinates and perhaps revised topology – that can be best achieved by a datum that links old databases with new by means of common identifiers.

Another argument is that the interoperability problem can be avoided by employing just one database across applications. There are several reasons why this is not practical. First, the functional requirements of a database are not universal. Data needs vary considerably between users – for example, there are situations where single-line representation of freeways is appropriate and preferable to dual-lines. Second, in a market-driven economy that is stimulated by competition, there is a danger in allowing any vendor, private or public, a monopoly over a critical information resource such as road centerlines. However, in the short term, the single-vendor approach may be most expedient.

5. Conclusions and prospects

The preceding sections have documented the types and magnitude of interoperability problems in location reporting. The data errors are presented not as a case of vendor malpractice, but as a consequence of an evolving state of art that is currently being overtaken by increasingly ambitious user requirements. As recently as 1980, development of street network databases was driven by demographic applications that required accuracy at no better than the general neighborhood or block face. Now just 20 years later, users are equipped with inexpensive GPS and portable computers; they seek navigation solutions, they demand carriageway resolution, 1–20 m accuracy and current information on attributes that change frequently (e.g. one-ways and turn restrictions, even construction and congestion). Emerging methodologies for emergency management and other applications are built around these heightened expectations. Inevitably, in the short term some of these requirements will not be met.

There are a number of initiatives under way, e.g. intelligent messaging and datums, to address these problems. One could argue that as the quality of databases evolves, the need for these solutions will diminish. On the other hand, it is reasonable to speculate that requirements will be even more stringent in the future, and that solutions will have to keep pace. One obvious area of development is improved resolution, from carriageway to lane. It is likely that navigational instructions will direct drivers to move into a certain lane in preparation for an approaching maneuver. The California Department of Transportation (Caltrans) is planning for a traffic management future where diversion of highway traffic is initiated several kilometres ahead of an incident, one lane at a time, to prevent concentration of such traffic into surrounding arterials.

These developments raise academic challenges. Data models must evolve to accommodate lanes (e.g. Fohl et al., 1996), and public initiatives to improve coordinate accuracy, such as datums, must anticipate the evolution of user requirements. New messaging methods have to be developed

for traffic information scenarios that not only report current status, but also describe previous or future events, such as the anticipated course of a toxic plume – some work on multi-dimensional data modeling is already under way, sponsored by the National Cooperative Highway Research Program (NCHRP) project 20–27(3). Based on the recent pace of development it is reasonable to believe that despite GPS, or perhaps because of it, the art and science of location reporting will remain an important area of endeavor over the coming decades.

Acknowledgements

References

Battelle Memorial Institute, 1997. Global positioning systems for personal travel surveys: Lexington area travel data collection test. Final Report. United States department of transportation, Federal Highway Administration, Office of Highway Management, Washington, DC.

Buttenfield, B.P., 1985. Treatment of the cartographic line. Cartographica 22 (2), 1–26.

Church, R., Curtin, K., Fohl, P., Funk, C., Goodchild, M., Kyriakidis, P., Noronha, V., 1998. Positional distortion in geographic data sets as a barrier to interoperation. Proceedings, ACSM Baltimore [document available at www.ncgia.ucsb.edu/vital].

Day, J.D., Zimmerman, H., 1983. The OSI reference model. Proceedings of the IEEE 71, 1334–1340.

Douglas, D.H., Poiker (formerly Peucker), T.K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. Canadian Cartographer 10, 112–122.

Fellegi, I.P., Sunter, A., 1969. A theory for record linkage. Journal of the American Statistical Association 64 (328), 1183–1210.

FGDC, 1998. Road Data Model – Content Standard and Implementation Guide (Working Draft). Federal Geographic Data Committee, Ground Transportation Subcommittee, Washington, DC.

Fletcher, D., 1999. Road data model workshop, Washington, DC. Meeting Notes.

Fohl, P., Curtin, K., Goodchild, M.F., Church, R.L., 1996. A non-planar, lane-based navigable data model for ITS. In: Kraak, M.J., Molenaar M. (Eds.), Proceedings of the Spatial Data Handling, Delft, 12–16 August, pp. 7B/17–29.

Funk, C., Curtin, K., Goodchild, M., Montello, D., Noronha, V., 1998. Formulation and test of a model of positional distortion fields. In: Proceedings of the Third International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Quebec City [URL www.ncgia.ucsb.edu/vital].

Goodchild, M., Hunter, G., 1997. A simple positional accuracy measure for linear features. International Journal of Geographical Information Systems 11 (3), 299–306.

Jaro, M., 1989. Advances in records linkage methodology as applied to matching the 1985 census of Tampa. Journal of the American Statistical Association 84 (406), 414–420.

Kim, K., Nitz, L., 1994. Application of automated records linkage software in traffic records analysis. Transportation Research Record 1467, 50–55.

Knuth, D.E., 1973. The art of computer programming. Sorting and Searching, vol. 3. Addison-Wesley, Reading, MA.

Mandelbrot, B.B., 1967. How long is the coast of Britain? Statistical self-similarity and fractional dimension. Science 156, 636–638.

Marx, R.W., 1986. The TIGER system: automating the geographic structure of the United States census. Government Publications Review 13, 181–201.

Noronha, V., 1999. Towards ITS map database interoperability – database error and rectification. Presented at International Workshop on GIS-T and ITS, Hong Kong. Forthcoming in GeoInformatica.

Noronha V., Goodchild M., Church R., Fohl P., 1999. Location expression standards for ITS – testing the LRMS cross streets profile (special issue on GIS Sharing and Standardization). Annals of Regional Science 33 (2) pp. 197–212.

Nystuen, J.D., Frank, A., Frank, L., 1997. Assessing topological similarity of spatial networks. In: Proceedings of the International Conference and Workshop on Interoperating Geographic Information Systems, Santa Barbara.

Quiroga, C.A., Bullock, D., 1998. Travel time studies with global positioning and geographic information systems: an integrated methodology. Transportation Research C 6, 101–127.

SAE, 1998. Surface Vehicle Information Report – Location Referencing Message Specification. Society of Automotive Engineers, Information Report J2374.

Siegel, D., Goodwin, C., Gordon, S.R., 1996. ITS Datum Final Design Report. United States Department of Transportation, FHWA Contract 61-94-Y-00001, Review Draft, 28 June 1996.

USDOT, 1997. The National ITS Architecture. A Framework for Integrated Transportation into the 21st Century. Federal Highway Administration, ITS Joint Program Office, Washington, DC. [URL www.odetics.com/itsarch].

USDOT, 1998. TEA-21 Critical Standards: Proposed Criteria and List of Critical Standards. United States Department of Transportation, ITS Joint Program Office, Washington, DC.

VITAL, 1999. The linear referencing profile – technical evaluation. United States Department of Transportation, FHWA Contract DTFH61-91-Y-30066, Final Report. [URL www.ncgia.ucsb.edu/vital].

Vonderohe, A.P., Chou, C.L., Sun, F., Adams, T., 1995. Results of a workshop on a generic data model for linear referencing systems. Proceedings, AASHTO Symposium on Geographic Information Systems in Transportation, Sparks NV.