Developing geographic information systems platforms for multijurisdictional transportation analyses: framework and techniques for spatial data sharing

Kai Han, Scott Minty, and Alan Clayton

Abstract: Geographic information systems (GISs) have been presented as a powerful analysing tool for civil engineers to help their decision-making processes. Building GIS platforms for transportation analysis involving multiple jurisdictions has been challenging, however, because of the complexity and difficulty associated with conducting data sharing and ensuring spatial data interoperability among GISs for transportation (GIS-T) data sets. In the context of western Canadian urban and rural areas, this paper investigates the issues related to GIS-T data sharing, establishes a conceptual framework, develops techniques supporting the framework by solving recurring data-sharing problems, and constructs a number of GIS-T platforms facilitating comprehensive multijurisdictional transportation analyses. In addition, based on the knowledge gained through solving real-world problems, the authors propose an open GIS-T platform consisting of a series of customized base maps, each being tailored to suit the needs of individual application and, as a whole, linked together by interoperability to better support transportation applications.

Key words: transportation engineering analysis, GIS, GIS-T, spatial data, interoperability, integration, data sharing.

1. Introduction

The past decade has seen widespread adoption of geographic information systems (GISs) as foundations for enterprise-wide data warehouses and as one of the major analysis tools capable of conducting spatial-related analyses in civil engineering fields such as environmental, water resources, and transportation.

Technologies such as the global positioning system (GPS) have made it easy to incorporate location information into the field data collection process. Civil engineers are taking advantage of this attached location information; for example, pavement condition data (international roughness index or IRI) were collected in conjunction with location data acquired by an on-board GPS receiver (EBA Engineering Consultants Ltd. 2001). These georeferenced data can help civil engineers to greatly improve their productivity through the use of GIS and related applications.

To reach the full potential of GIS and related data sets, data sharing is needed. In both public and private sectors, transportation organizations produce and maintain large GIS data sets as a part of their information systems. These data exist as individual data sets with very limited data-sharing capability. When transportation analysis or planning involves

Résumé : Les systèmes d’information géographique (SIG) ont été présentés comme étant un outil d’analyse puissant pour les ingénieurs civils, un outil qui peut les aider dans leurs processus de prise de décision. Cependant, la construction de plate-formes SIG pour l’analyse des transports impliquant de multiples territoires de compétence présente un gros défi, en raison de la complexité et de la difficulté à partager les données et à s’assurer de l’interopérabilité des données à référence spatiale parmi les ensembles de données SIG pour les transports (SIG-T). Dans le contexte des régions urbaines et rurales de l’Ouest canadien, cet article étudie les questions ayant trait au partage des données SIG-T, établit un cadre conceptuel, développe des techniques supportant ce cadre en résolvant les problèmes récurrents de partage de données, et construit un nombre de plate-formes SIG-T facilitant les analyses complètes des transports impliquant de multiples territoires de compétence. De plus, en se basant sur les connaissances acquises par la résolution de vrais problèmes, les auteurs proposent une plate-forme SIG-T ouverte consistant en une série de cartes de base sur mesure, chacune étant adaptée pour répondre aux besoins d’une application individuelle et, dans l’ensemble, reliées ensemble par interoperabilité afin de mieux soutenir les applications dans le domaine des transports.

Mots clés : analyse en ingénierie des transports, SIG, SIG-T, données à référence spatiale, interopérabilité, intégration, partage des données.

[Traduit par la Rédaction]
multiple jurisdictions, GIS data sharing and integration become inevitable. Sharing of GIS data is a topic receiving increasing interest in the current GIS for transportation (GIS-T) community.

When attempting to integrate and share data from different GIS-Ts, many issues arise. Generally, GIS-Ts are developed as individual systems with jurisdictional scopes and application-specific considerations, using combinations of different software, hardware, spatial data sets, and representation schemes. These differences between GIS-T platforms can lead to potential interoperability problems, which impede GIS data sharing and multijurisdictional analysis.

To achieve the goal of building of multijurisdictional transportation GIS-T platforms, the authors investigate common issues related to GIS-T data sharing and spatial data integration and establish a conceptual framework and develop associated techniques based on many years of data-sharing experience. The application of the techniques under the guidance of the framework culminated in the development of a number of regional GIS-T platforms that facilitate the transportation analyses and decision making.

1.1. Geographic information systems for transportation data sharing and interoperability

The spatial and temporal dimensions of transportation systems make GIS particularly suitable for analysing transportation problems. Transportation analyses often involve multimodal, multisectoral, multiproblem, and multidisciplinary considerations focusing on the interaction between the transportation and activity systems of a region (Manheim 1979). Realizing the advantages of GIS, transportation was one of the earliest application fields of GIS (George 1991).

To accommodate various types of transportation applications, GIS-T professionals have developed a multitude of GIS-T platforms. In most cases, the development of GIS spatial and attribute data is an expensive, time-consuming, and labour-intensive process. The utilization of these valuable data sets and many legacy databases already developed by civil engineers requires data sharing for regional, national, and international analysis involving multiple jurisdictions and transport modes, since transportation networks and flow do not stop at jurisdictional borders. In addition, the value of existing GIS-T information can be multiplied through data sharing among agencies and departments. Sharing of GIS-T data helps link geospatial data islands, save time and money on data acquisition, and most importantly, conduct comprehensive transportation analysis. A recent example of this type of analysis is the project sponsored by Transport Canada to study the options for international harmonization of vehicle weights and dimensions (VW&D) under the North American Free Trade Agreement (NAFTA) in the western region, which encompasses the provinces of Manitoba, Saskatchewan, Alberta, and British Columbia and the states of Minnesota, North Dakota, Montana, Idaho, and Washington. Figure 1 shows the regional highway network integrated to support the study (Clayton et al. 2002).

The full potential of data sharing can be achieved only through interoperability. According to the OpenGIS guide (OGC 1998), interoperability is defined as the ability of a system or components of a system to provide information portability and interapplication, cooperative process control. Interoperability, in the context of the OpenGIS specification, is software components operating reciprocally (working with each other) to overcome tedious batch conversion tasks, import-export obstacles, and distributed resource access barriers imposed by heterogeneous processing environments and heterogeneous data (OGC 1998). Geographic information systems for transportation applications have some unique characteristics, which makes GIS-T an interesting and demanding field for implementing interoperability. For example, GIS-T applications such as networking, routing, and linear referencing make strict and diverse demands on GIS spatial data. To accommodate these applications, spatial data structures of GIS-T systems vary significantly. In the GIS-T field, many data-sharing issues are related to spatial data. Therefore, it is important to handle interoperability issues from the spatial data level to accomplish data sharing among GIS-T systems.

In the past 7 years, the authors have worked on a series of contracted research projects requiring integrations of transportation-related spatial and attribute data in the whole western Canada region. It was this experience that gave rise to the idea of this paper. Recent principal examples of the projects are as follows:

1) The integration and transformation of three versions of the Manitoba provincial highway base map. Each version possesses different projections, scales, and spatial data structures (Han 2001a).

2) The integration of provincial and state highway GIS networks in western Canada and the adjacent United States, culminating in a highway network that contains four provinces (Manitoba, Saskatchewan, Alberta, British Columbia) and five states (Minnesota, North Dakota, Montana, Idaho, Washington). Key differences among the involving data sets are different scales, multiple software packages, a variety of linear-referencing (LR) schemes, and distinctive spatial data structures (Clayton et al. 2002).

3) The construction and integration of railway maps in western Canada, with additional compatibility with the highway network and aerial photography (Minty et al. 2002).

4) The application of GIS in the analysis of heavy truck accidents in the Prairie Provinces (Manitoba, Saskatchewan, and Alberta). The investigation involves integrating highway inventory databases, traffic count and vehicle classification data, and historical accident databases with the regional highway network to identify spatial and temporal characteristics of heavy truck accidents (Montufar 2002).

1.2. Recurring interoperability issues regarding spatial data

Many issues arise when sharing data among GIS-Ts is attempted. Based on the authors’ experience, this paper gives special attention to the following recurring interoperability issues related to spatial data: currency and completeness, accuracy, level of generalization, spatial data structure difference, and metadata for spatial data.

1.2.1. Currency and completeness

Transportation networks are continually changing; new fa-
cilities are built, existing roads may be rerouted, and a railway may be realigned. Spatial data in a GIS should reflect these changes in a timely fashion. Otherwise, the information provided by that specific GIS and the analysis based on it will soon be rendered incomplete and incorrect.

Currently, most GISs lack the flexibility of updating their spatial data promptly, since they rely on a single source to build and update spatial data. Updating and improving spatial data quality can be difficult and are often left undone until new GISs are developed. A long updating period for spatial data hinders data sharing. Mislinking and mismatching can occur during the data-sharing process when two data sets are collected at different times.

1.2.2. Accuracy

Two types of accuracy are involved in the data-sharing process: absolute accuracy and relative accuracy. Interoperability is more affected by relative accuracy. To ensure correct spatial analysis, good relative accuracy is critical. Poor relative accuracy causes large discrepancies between spatial data sets, which can lead to inaccurate and inefficient linking-matching and consequently analysing results. Since relative accuracy is often interrelated with absolute accuracy, better accuracy is desired in general for conducting GIS data sharing.

The accuracy of a specific GIS spatial data set is often limited by the data sources available at a certain time and within a certain financial restriction. Computational ability is another factor affecting spatial data accuracy. New technologies provide increasingly accurate spatial data, which not only helps provide better representation of the real world, but also makes possible the analysis involving data from different sources.

1.2.3. Level of generalization

All maps and spatial data sets are simplified models of reality (Heywood et al. 1998). Generalization refers to the mapping process in which complicated real-world features are simplified for clarity. In conjunction with thematic symbols and labels, the generalization process helps to convey the information clearly. Depending on the purpose of the GIS and the scale of the base maps, the levels of generalization can vary greatly from one GIS to another. For example, a city can be represented as a point in one GIS, an area in another, and thousands of links and nodes in a third.

When the same real-world features are represented differently in various GISs, data sharing among them becomes difficult. The issue is further complicated by the fact that the level of generalization is often intertwined with scale, accuracy, and level of detail. Highly generalized spatial data are more likely to have a smaller scale, lower accuracy, and less detailed spatial data features. On the other hand, fewer generalized spatial data are closer to reality, containing more detail and usually associated with better accuracy resulting from large-scale mapping.

Differences in levels of generalization and consequently in the spatial data structures make matching-linking these data difficult. Therefore, matching spatial data using the buffering technique and cross-layer linking is often required to identify the relationships between spatial data features from different data sets.

1.2.4. Spatial data structure difference

Spatial data in GISs are structured to represent the real-world phenomenon that they are designed to tackle and often restructured during their use to accommodate certain applications. The structure of the data storage in a specific GIS application may impede data sharing with other GISs and the adoption of newer, more accurate spatial data. Spatial data from other sources and newly available spatial data may not be compatible with the original data structure. This incompatibility makes it difficult to conduct direct conversions. A good example of this is the difference in spatial data structures for road networks. Highways might be seg-
mented from one highway intersection to another in a provincial-level network, leaving local accesses not represented. On the other hand, the same highway going through an urban area was typically segmented at every intersection, no matter how big or small (e.g., back lane), to form a local street network. Another example, shown in Fig. 2, illustrates that different spatial data structures (single centreline versus double centreline) between Manitoba and North Dakota highway base maps create difficulties in making an integrated network for transportation applications such as vehicle routing and linear referencing.

1.2.5. Metadata for spatial data

Metadata are data about data (Microsoft Corporation 1998). They give the GIS user information about how to interpret and handle the spatial data of concern. A complete set of metadata should include information about identification, data quality, spatial data organization, spatial reference, entity and attribute, distribution, reference, citation, time period, and contact.

The U.S. Federal Geographic Data Committee (FGDC 2001) has created a standard for digital geospatial metadata. It provides a common set of terminology and definitions for the documentation of digital geospatial data. This standard has not been widely adopted, however, and very few spatial data sets are provided together with standard metadata. Also, GIS software has very limited support for metadata. The lack of metadata renders many spatial data sets difficult to integrate with other spatial data.

2. Geographic information systems for transportation data-sharing endeavours and research approach

Data sharing and interoperability between GISs are issues of increasing concern in GIS research. Dedicated conferences were held in 1997 and 1999, interested parties formed the Open GIS Consortium, and there have been numerous efforts to promote interoperability among GISs to overcome problems concerning GIS software, computer platforms, data format, and spatial data structure issues (Beskpopko et al. 1997; National Center for Geographic Information and Analysis 1997; Noronha 1997; Xiong et al. 1999; Dueker and Butler 2000; Han et al. 2000; OGC 2002; Minty et al. 2002). The GIS-T community is establishing a common data model and linear-referencing system to enable spatial data sharing (National Cooperative Highway Research Program (NCHRP) 20-27) (Goodwin 1996; Adams et al. 2001). Data-sharing efforts were also given in the area of integrating GPS data into GISs (Czerniak 2002). The most recent attempts include the drafting of the road component of the Data Content Standards for the Geographic Information Framework sponsored by the U.S. Bureau of Transportation Statistics (BTS) and the development of the Canadian National Road Network (NRN) standard by Natural Resources Canada (2002).

Realizing the general goal of this research is to build multijurisdictional GIS platforms for transportation analysis, the authors understand that GIS-T data from various data sources need to be integrated to form the platforms. These data sets may differ in coverage, accuracy, structure, and generalization. The essence of the processes developed in this paper is to bring the spatial data to attribute data, rather than ask attribute data to come to the maps, namely, gather GIS-T data from various sources, transform the spatial data structure of each data set to accommodate the final application, and merge these manipulated data sets. For example, when a regional highway network for vehicle routing purposes is being established, simple single-centreline base maps are usually preferred. Therefore, to make the spatial data structure of each data set compatible, special techniques were developed to convert multicentreline presentation in certain jurisdictions to single-centreline presentation and consequently merge the data sets with proper attribute data attached.

To achieve this flexibility in transforming spatial data structures, the authors developed a series of specialized techniques in our research to modify spatial data structure and ensure interoperability between data sets and established a conceptual framework to provide a skeletal structure in which GIS users can conduct the linking processes on spatial data sets. Under the framework, linkages between GIS-T spatial data are established by applying the appropriate techniques. Consequently, through spatial data linkage, attribute data can also be rendered interoperable.

3. The framework

The linking and matching of GIS-T spatial data is a process that involves several phases and employs several tech-
niques. The steps to follow vary from case to case. To incorporate these various components and provide a data processing scheme and data structure blueprints when new spatial data sets are to be processed, a framework was developed during the research.

To simplify the data-sharing process, the framework handles spatial data and attribute data separately. Since the spatial data, by definition, have common location information that can naturally be used as a linking agent, the framework is constructed around, and has its emphasis on, spatial data. When spatial data sets are linked, GIS users can then use the interoperable spatial data as bridges to allow the free flow of attribute information.

3.1. Framework structure

The framework developed in the research forms a pyramid shape in accordance with the generally accepted view of the scaling factor in mapping and spatial data generalization. The framework is illustrated in Fig. 3.

The framework consists of layers representing different levels of abstraction of the real world. These layers are ordered by their levels of abstraction. Naturally, the level of abstraction is correlated with mapping generalization and scale. The lower the level in the framework, the larger the scale, and the closer spatial data get to the real world. Consequently, to render the details of the real world, spatial data become more complicated and usually have better accuracy. Conversely, fewer spatial data are necessary as the abstraction level goes up in the framework. In general, the result is that, even through the real-world entities they represent are the same, no two spatial data sets are identical.

Depending on the scale of the spatial data set, spatial data are structured following certain generalization rules. For example, a city can be represented as a myriad of points and lines depicting a dense urban street network in the city transit planning data set and merely a point in the world city data set. To identify these abrupt changes in the generalization process, spatial data layers are further grouped into tiers. Within a tier, the spatial data representing same real-world entities do not change type from one spatial data layer to another. This hierarchical structure of the framework facilitates the selection of appropriate techniques to be used to render different spatial data sets interoperable. The number of layers (or tiers) in the framework is not fixed. It varies depending on how many different spatial data sources are used in the data-sharing project.

To focus the data-sharing effort, the conceptual framework has a limited scope in terms of spatial coverage, usually covering the study region, which can vary from case to case. Only spatial data sets (or portions of them) that are inside the study region are to be handled by the framework.

3.2. Dimensions

There are multiple dimensions in the framework. Besides the aforementioned scaling – level-of-abstraction and horizontal layer – tier-type dimensions, the framework also provides the following dimensions:

1) User level dimension — GIS-T users can range from sand truck drivers, who determine when, where, and how much sand to spread, to policy makers, who plan long-term multimodal transportation systems. These different users have different requirements of spatial data. Planning-level user groups often require a larger scope of spatial coverage and are concerned with systematic interaction more than specific details. On the other hand, user groups coping with local issues, such as civil engineers working on a specific project, need data accuracy and a greater level of detail to support their analyses. Therefore, there is a correlation between the scaling – level-of-abstraction factor and the user group level. The framework shows the interconnections of the two perspectives.
(2) Vertical view dimensions — One vertical view is top-down. With this viewpoint, GIS-T users approach data sharing from a higher level. For example, a provincial transportation engineer can fit lower level spatial data features from a detailed municipal road GIS into the provincial GIS to share information residing on the lower level GIS. Conversely, the bottom-up view of the framework assists data-sharing efforts by allowing lower level users to take advantage of higher level layer information.

(3) Time dimension — Since transportation systems are dynamic, systems structures and components change over time. Usually it takes time for a GIS-T to reflect real-world changes. The time from the actual change taking place to the change being perceived by the GIS-T varies from one data set to another. To ensure the correctness of analyses, an appropriate time-stamping process should be applied to the data set. When data sharing is concerned, synchronization between GIS-T data sources is essential.

3.3. Theoretical foundations

Several important characteristics of spatial data and fundamental relationships among data sets form the theoretical foundation of the framework and the associated techniques.

In the course of this research, location information has the basic role as the linking agent. Although spatial data sets from different sources differ in accuracy, data structure, and generalization, the location information that is inherent to spatial data provides the foundation for the framework to identify spatial relationships among different spatial data sets as long as they represent the same real-world entities. Based on these spatial relationships, matching and linking of the spatial data can be conducted effectively.

The spatial relationship defines the way in which two spatial data features (i.e., points, lines, and polygons) interact with each other. Common examples include “containing”, “intersecting”, and “within certain distance”. They can be further broken down to different categories by spatial data feature types. It is also important to note that the actual matching–linking process is conducted in a cross-layer fashion, which differs from ordinary within-one-layer topological analyses.

The data-sharing processes of the framework explore the spatial relationship in different ways. These processes can be grouped into two types (explicit and implicit), in terms of their approaches.

Explicit approaches directly investigate spatial relationships and derive matching–linking relationships based on spatial proximity. Generally, spatial data from different sources are in the vicinity of the absolute location of the real-world entity they represent. The spatial proximity can be explored using the buffering technique, which creates a buffer zone around one particular spatial feature (i.e., a point, line, or polygon) and finds common spatial data features inside this buffer zone. This spatial proximity consequently renders matching–linking relationships (Han et al. 2000).

Implicit approaches employ advanced techniques such as linear referencing and centreline derivation to extend the scope of spatial data sharing. In these approaches, spatial relationships are used internally under the cover of mechanical processes. For example, the centreline-derivation technique finds points having equal distances from both sides of the roadway along a specific route, and the linear-referencing technique converts one-dimensional tabular data with certain distance marks, e.g., milepost, into two-dimensional (sometimes three-dimensional) spatial data along the line set. With these approaches, spatial data can be restructured and rendered in different ways to fit the needs of different applications, and new opportunities are also created for information residing on legacy databases that can be converted into a GIS format compatible with newly developed GISs.

3.4. Data sources

The framework can handle a broad spectrum of GIS data. Through the application of appropriate methods and techniques, these georeferenced GIS data can be blended into the framework and have spatial data features linked or matched.

The information sources are not limited to existing GIS data sets. Any location-referenced information can also be incorporated into the framework. These additional information sources will enhance the spatial data and attribute data aspects in the framework. These information sources can include (i) basic land survey systems, (ii) geodetic control points, (iii) engineering drawings and computer-aided design (CAD) drawings, (iv) aerial photography and satellite imagery, (v) GPS field surveys, and (vi) mathematically calculated grids.

3.5. Outcomes

The framework helps to achieve better interoperability, which benefits spatial data sharing and, consequently, attribute data sharing among different GISs. The results of the matching–linking of different GIS data sets are generally delivered in the form of a translation table. With the table, further data sharing is made possible on the basis of this foundational relationship.

An important output of the framework is that it provides an avenue to inexpensive, better quality, and more compatible spatial data sets for every data source involved in the framework. With interoperability at the spatial data level, spatial data sets can swap spatial data features interchangeably among different data sources, taking advantage of the best (most accurate, current, complete, and appropriately generalized) spatial data available. The process can be accomplished through the application of techniques associated with the framework. Improved spatial data quality benefits GIS analysis and further data-sharing efforts. This process also creates new opportunities for existing GIS-Ts to enhance and update their spatial data, which will better reflect the real world and facilitate future application development.

The output spatial data sets can be more accurate, more complete, constantly updated, and, most importantly, completely compatible with the original GIS-T application. Therefore, GIS-T applications do not need to make any changes to accommodate new spatial data, and the transition process will be minimized.

4. Technique development

To conduct data sharing accurately and efficiently, the
framework must be supported by a group of specialized techniques. These techniques can be applied to spatial data layers introduced into the framework to help link or match spatial data sets, manipulate spatial data structures, integrate spatial data sets, and populate the attribute data table. Although these techniques are developed to solve individual problems, they can be customized to handle more generic tasks in future processes. Some of the techniques employ existing functions found in off-the-shelf GIS software. However, most of the techniques involve in-house computer programs developed to solve specific problems. The accumulation of these techniques forms a growing toolbox for data sharing. Subsequently, the data-sharing process can take advantage of previous endeavours to save time and maintain the quality of the results.

A series of techniques were developed to solve interoperability problems encountered during the research. As discussed previously, the techniques focus on spatial data processing. The techniques are grouped into two categories, based on their purposes: (i) spatial data restructuring and enhancement, which involves manipulating the spatial data structure to meet specific needs, and (ii) spatial data linking and matching, which uses location proximity to identify relationships among spatial data sets.

To help assign attribute data to their matching spatial data, a special group of techniques was also developed for attribute data integration.

4.1. Design considerations

Several technical aspects were considered during the development of the techniques. They include the balance between automation and human intervention, the flexibility to handle different data structures, the employment of existing GIS functions, the use of common file formats to ensure accessibility to various data sources, and the quality-control process to identify anomalies and remove ambiguities.

Since spatial data sets often have complicated structures that necessitate extensive numerical calculations when data-sharing techniques are implemented, it is desirable to implement the techniques in an automated fashion. User intervention is still required to handle extraordinary situations, however. For example, in the process of linear referencing, when tabular data have mileposts lying outside the highway control section range, user intervention is needed to determine the locations based on engineering judgments. Techniques developed in this paper involve combinations of automated processes and user intervention.

To accommodate different GIS spatial data sets, the authors have developed programs with the flexibility to handle different data structures. This flexibility is achieved through altering the parameters of the program. This practice also provides reusability of the code to fit different types of problem solving.

Whenever possible, the techniques developed in the paper employ functions provided by commonly available GIS software to eliminate unnecessary programming. Software for GISs provides general-purpose functions and often comes with an internal programming language, which can be used to customize functions and automate work processes. However, it is often found inadequate to perform data sharing in many cases. When GIS software cannot handle the situation (e.g., deriving theoretical centrelines from double centrelines), specialized in-house programs are developed to implement the technique.

To provide the highest degree of accessibility to spatial data sets, the authors have adopted the text format as the common spatial file format internal to the framework. Similarly, the dBase format is adopted as the common file format for attribute data. Using these common formats facilitates data-sharing processes involving existing GIS functions and in-house programs. In addition, commonly used proprietary GIS data formats are supported during data input and output processes to facilitate end usage.

In the research, in-house programs are developed to solve data-sharing problems efficiently. A number of factors are considered during the programming. Among these factors, “performance” is a major concern, especially when large amounts of spatial data are to be processed. Non-crucial programming features such as graphical user interfaces are eliminated from the in-house programs to simplify the program and reduce running time. “Accuracy” is another factor. To ensure accuracy of the results, the in-house programs often require time-consuming floating-point calculations. To address the large computation needs with floating-point calculations, C++ programming language is used for implementing the techniques to achieve better performance (i.e., faster calculation) without sacrificing the accuracy of the calculations.

The results of the matching–linking processes are contained in multiple tables. Before undertaking any further data-sharing efforts, it is necessary to remove ambiguity within the tables. Such a cleaning process uses database queries to search records, identify anomalies, and spot ambiguities. Then a combination of automatic and interactive processes is applied to resolve the problems. Using the resulting match–link table, data sharing can be conducted without introducing ambiguity throughout the framework.

4.2. Development examples

There is a series of techniques developed by the authors to handle different data-sharing problems. These techniques are important in helping bring the spatial data aspect of GIS to attribute data or vice versa. Details of these programs are documented by Han (2001a). Two of the key techniques are presented here to provide a closer look at the typical technique development process.

4.2.1. Automated theoretical centreline derivation procedure

This technique is developed to derive a theoretical single centreline from double centrelines that are used to represent divided highways in some GIS-T base maps. It allows GIS users to move from the double-centreline representation to the single-centreline representation, or vice versa, and facilitates two-way information transformation while ensuring data integrity.

Single- and double-centreline representations are two commonly used schemes in the current GIS-T field to represent divided highways (Han 2001b):

1. Single-centreline representation using one line feature depicting the centreline of the total roadway to represent a divided highway — When this scheme is carried
through the entire highway network, all highway sections are represented by single centrelines, whether they are divided or undivided. The internal compatibility and simplicity of this scheme facilitates transportation planning, network analysis, and small-scale presentation of the highway system.

(2) Double-centreline representation in which divided highways are portrayed by two centrelines following the centrelines of the two separated carriageways — This representation, often associated with ramp details at interchanges, provides a more detailed picture of the divided highway infrastructure than single-centreline representation. Although this scheme facilitates applications like asset management processes, representation, and large-scale presentation (e.g., traffic data or pavement condition data (IRI)), many network-based applications are not well supported by the double-centreline scheme.

The advantages and disadvantages associated with single- and double-centreline representation schemes are summarized in Table 1.

AWARE of the advantages and disadvantages of the two representation schemes, transportation agencies often have difficulty choosing schemes. Decisions can depend on many factors, e.g., data source, main applications of the system, and cost. For example, Manitoba Transportation and Government Services has developed three GIS-T base maps over the past 7 years, two of which use the single-centreline scheme (the first- and second-generation base map), and the latest of which uses the double-centreline scheme (the third-generation base map).

For a rural highway network, such as most roads in the western Canada region, it is generally accepted that single-centreline representation has many advantages over double-centreline representation for its simplicity, networkability (one intersection versus four intersections when two divided roads intersect), and interlinking ability (easy to link GPS points and plots accidents by direction). In certain cases, such as a dual divided freeway or a divided toll road built on two sides of the existing freeway, double centrelines can be useful, provided that each of the two outer centrelines is treated appropriately as one-way streets.

To take advantage of both representation schemes, interoperability between the two types of base maps is highly desirable. Although line matching can be established between a single-centreline base map and a double-centreline base map using buffering techniques, the efficiency of the process and the accuracy of the result are not always satisfactory when the two base maps are of different scales.

Since double-centreline base maps are generally newer and have better spatial accuracy and a higher level of detail than their single-centreline counterparts, it is desirable to derive new single centrelines from the most current double-centreline base maps in an efficient manner to ensure better spatial data accuracy.

The technique developed in the research employs several in-house programs and calculates the theoretical centreline from existing double centrelines in the base map. Figure 4 shows an example of the derivation. Figure 4a shows the double centrelines of the divided roadways of TransCanada Highway control section 1001290 in Manitoba east of Winnipeg, and Fig. 4b shows the derived theoretical centreline tracking in the middle of the original double centrelines.

4.2.2. Topology standardization procedure for complex networks

Most urban centres maintain street centreline (SCL) GIS systems to represent their street networks. Various transportation-related information can be attached to these SCL GIS systems. Examples include functional classification, accident data, and pavement-condition data.

During data-linking processes, a topology issue arises from time to time. Owing to the lack of uniformity in the direction of line entities in the SCL spatial data set, directional traffic counting data cannot be assigned to the SCL automatically. Since many potentially linkable transportation data are directionally sensitive, and the amount of directional data often makes manual assignment prohibitive, it is important to ensure standardized topological directions in the SCL network.

In many cases, the original topology reflects the direction of digitization and is often haphazard. To accommodate large amounts of directional data, topological directions of the network links need to be modified and standardized according to a database convention. Once the standardized top-
ological direction is established, the base map can easily receive traffic counting data stored in a variety of databases through automated processes.

To overcome this problem, the authors developed a technique and associated in-house computer programs to reverse the topological directions of part of the line features in the spatial data, giving standardized direction to the line features while maintaining the integrity of the original data set. For example, in the application in the City of Winnipeg, the design criterion for restructuring was to standardize all topology to run mostly west to east and south to north. These are the directions compatible with the adjacent provincial highway control section system. Figure 5 shows the standardization process from the original directions to the standardized directions of the streets in the City of Winnipeg.

The essence of this technique is to present the GIS-T platform in a logical way to attribute databases to help bring directional data to a GIS environment without incurring any change to the legacy database.

5. Developing integrated platforms for geographic information systems for transportation

By implementing the appropriate techniques following the scheme provided by the framework, GIS-T platforms for multijurisdictional transportation analyses are developed during the research. The data-sharing processes are under the guidance of the framework developed in the research and conducted in the following manner: (i) GIS-T data sets from different information sources are separated into spatial data sets and attribute data sets; (ii) the spatial data sets are inserted into the framework according to their levels of abstraction; (iii) GIS-T users decide which way the shared information should flow (upward, downward, or both); (iv) according to the direction of the information flow and relative position of the spatial data layers (or tiers), suitable methods are chosen and relevant techniques are developed—applied to the spatial data sets; (v) the resulting linking–matching tables are used to help move attached attribute information across spatial data layers; (vi) to take the result of the framework a step farther, the GIS-T user can derive a spatial data set based on the best quality data sources available with a suitable level of abstraction to enhance the quality of their particular GIS-T application, thus facilitating future data-sharing projects.

The following sections present two examples of integrated GIS-T platforms.

5.1. Platform for Manitoba Capital Region

The authors integrated the Manitoba Provincial Highway base map, the Natural Resources Canada map fabric, and the City of Winnipeg street centreline (SCL) map to create a GIS platform for the new Manitoba Capital Region Transportation Planning Model (Clayton et al. 2001).

The road network has different densities in urban and rural areas. For the purpose of supporting a regional transportation planning model, the network connectivity at urban and rural interfaces and setting up data structures that can help correlate spatial features among separate networks are emphasized. Figure 6 shows the road portion of the platform and a detailed example of the integration at jurisdictional boundaries.

5.2. Platform for the prairie region winter weight premiums and spring weight restrictions study

This regional platform integrated GIS base maps from Manitoba, Saskatchewan, Alberta, Minnesota, North Dakota, Montana, the cities of Winnipeg, Edmonton, and Regina, © 2003 NRC Canada
three rural municipalities, and two national parks in the prairie region. The winter weight premiums (WWP) and spring weight restrictions (SWR) through the study year are attached to the platform using various techniques (e.g., linear referencing) to provide a regional perspective of this transportation regulation issue. Analyses of the rationalization and harmonization options of WWP and SWR were further conducted based on the platform to facilitate regional truck transportation (Montufar et al. 2000). Figure 7 gives an example of SWR affecting the regional highway network during the spring thaw period.

5.3. Open platform for geographic information systems for transportation

Since transportation applications have evolved in so many directions, the requirements on GIS-T base maps have diversified greatly. There is a consensus that “no one GIS base map can serve all applications”. Base maps for GIS-T need to be application specific (Minty et al. 2002). At the same time, computer, mapping, remote-sensing, and GIS technologies have quickly progressed to provide more accurate, detailed, and up-to-date spatial data.

To take advantage of these new spatial data and provide better support to various applications, the authors propose the development of an open GIS-T platform. As a natural extension of the framework developed in the paper, this platform incorporates multiple data sources and provides interoperability through manipulating the spatial data structure and integrating spatial and attribute data by employing various techniques. As more and more GIS data sets are included in the framework, a repository of data containing original spatial features from different sources (with their sources identified and having the best accuracy and currency) can be created. Based on this repository, an interoperable GIS-T platform can be derived and maintained at a high standard, which in turn facilitates future GIS-T interoperability.

The strength of the platform lies in the fact that it consists of a series of customized base maps, each being constructed through blending, melding, and customizing to suit the needs of individual applications. Meanwhile, the interoperability offered by the platform ensures linking capability among individual base maps and allows for exploitation of the full potential of GIS-T applications using diverse spatial entities in terms of scale, schemas, and accuracy. The advantages of this solution include (i) facilitating access to and use of more information sources, (ii) providing a more transportation application friendly spatial data structure, and (iii) allowing for continuous updating and enhancement of the spatial data set to give a more accurate picture of the real world.

The platform can also support new GIS-T developments by providing spatial data sets that have the highest quality data available and built-in interoperability with current GIS-T. Ideally, a GIS-T designed for multiple purposes should consist of two (or more) sets of spatial data specifically structured to meet individual needs, such as conceptual analysis, detailed asset management, and visual presentation. The requirements of different purposes can be met for the most part without compromise. Interoperability among these spatial data sets binds them together into an integral platform for transportation applications.

Compared with other proposed methods designed to facilitate data sharing and interoperability, such as NRN and BTS data content standards, the open platform takes a different approach. Instead of compiling a single set of spatial data and recommending existing systems to migrate to this common set of representations, the open platform gives transportation application users (i.e., civil engineers) the freedom to choose the most suitable spatial data sets in terms of level of generalization, spatial data structure, and attribute data attachment. This approach lets GIS-T users take full advantage of the existing GIS data sets without compromising data integrity or incurring the cost of time-consuming GIS migration processes. In the meantime, GIS-T users can enjoy the additional benefit of frequent updating and incorporation of the newly available data.

The open platform also offers the potential of strengthening the GIS data standard through the incorporation of the proven concept and procedures into the standard development process.
6. Conclusion and recommendation

Geographic information systems for transportation data sharing and interoperability are important issues for civil engineers conducting transportation analyses involving multiple jurisdictions. With data sharing, GIS-T users can greatly enhance their analyzing capability and productivity. Interoperability also makes possible the sharing of valuable information among spatial data sets without sacrificing data integrity. This flexibility gives GIS-T users the freedom to migrate to a GIS-T platform that can be constantly enhanced throughout its life cycle.

Based on the experience acquired through research projects conducted in the western Canada region, the authors established a new conceptual framework, developed techniques supporting the framework by solving recurring spatial data-sharing problems, and constructed a number of GIS-T platforms to facilitate comprehensive multijurisdictional transportation analyses and decision-making.

The potential benefits of the framework and techniques developed in this paper can be extended to the whole GIS-T community with the integration of the framework and techniques into GIS data standards and common GIS software.

The diversity of GIS-T applications prevents the use of a common set of spatial data for different applications. The authors propose an open GIS-T platform as a solution to provide customized spatial data sets to multiple GIS-T applications while maintaining interoperability among data sets. This will allow more applications to share spatial data through the interoperable platform.

Acknowledgements

The authors would like to thank Transport Canada; Natural Resources Canada; the Prairie Farm Rehabilitation Administration (PFRA); British Columbia Ministry of Transportation; Alberta Infrastructure; Saskatchewan Department of Highways and Transportation; Manitoba Transportation and Government Services; Manitoba Conservation; the transportation departments of Washington, Montana, North Dakota, and Minnesota; the cities of Edmonton, Saskatoon, Regina, and Winnipeg; EBA Engineering Consultants Ltd.; and Atlis Geomatics Inc. for providing GIS data and related information to facilitate the research reported in this paper.

References


