SPATIALLY DISTRIBUTED NAVIGABLE DATABASES FOR INTELLIGENT VEHICLE HIGHWAY SYSTEMS

Thomas J. Cova and Michael F. Goodchild
NCGIA/Department of Geography
University of California Santa Barbara
Santa Barbara, CA 93106
cova@geog.ucsb.edu, good@geog.ucsb.edu

ABSTRACT

Intelligent Vehicle Highway Systems (IVHS) are placing formidable challenges before the GIS community regarding the issues of dynamic data collection and sharing. This paper outlines the problem and reviews the requirements that are driving the need for a distributed spatial database in meeting the needs of IVHS. A set of contemporary GIS architectures is described and evaluated regarding these requirements.

INTRODUCTION

In any major U.S. city, data regarding its road network is managed at a number of geographically separate sites by a wide variety of historically separate agencies. A single city road segment may fall within the political jurisdiction of any number of entities, most obvious of these being its host city, county, and state. This hierarchical view of space can be further divided into transportation districts, school districts, and fire districts, to name just a few partitionings in the public sector. All of these entities have different conceptual views of the road network and store varying information regarding its state and condition. In this way, they all share an interest in maintaining a database regarding the roads under their jurisdiction. Historically, these autonomous organizations have maintained separate databases, but IVHS is narrowing these spatial information chasms by strongly inducing new forms of data-sharing.

IVHS aims to reduce traffic congestion by computerizing cars, their routes, and the roads themselves. This computerization is intended to decrease transport problems and environmental pollution through reduced fuel consumption and less accidents. Example areas for automation include route guidance, toll collection, hazard warnings, fleet dispatch, commercial services directory, collision avoidance, and driving with full-time automated control.

A database capable of supporting the requirements of these systems will inevitably demand a high level of cooperation between many disparate sources of spatial data. In this way, IVHS depends on the political and technical ability of a diverse community to seamlessly integrate geographic data from myriad sources into a unified resource. This is not to say that the database need necessarily be physically centralized, but rather, that users require a unified resource regardless of how the data may be distributed geographically. This problem of "data fusion" was identified early within the context of IVHS and continues to receive significant attention (Lee, et. al., 1989; Summer, 1991; Schiff, 1993).

Real-time data fusion points to an act of dynamic data-sharing and cooperation that has yet to be fully realized in the spatial database community. It stands as a direct challenge to existing methods for rapidly collecting, managing, and disseminating accurate spatial data. A technical side of this challenge was recently presented to the wider database management community within the context of vehicle navigation (Egenhofer, 1993), and approaches aimed at improving our ability to manage real-time, distributed, geographic databases for IVHS are surfacing (Choy, Kwan, & Leong, 1993). In short, IVHS is encouraging the fusion of many cross-site road database seams, and the future points towards a transparently-distributed, navigable database that is managed in real-time by disparate parties through an emerging spatial database cooperative.

The first part of this paper briefly describes navigable data models and distributed databases. The following section presents a set of IVHS requirements that drive the need for a navigable database cooperative. The final section describes a classification of contemporary GIS architectures and evaluates the potential for these architectures to support such a cooperative.

BACKGROUND

Navigable Data Models

The fundamental real-world element to model in supporting IVHS is the transportation network. A primary approach to modeling a road network involves a vector model of space, where nodes represent road intersections and links represent road segments. Connectivity between elemental objects in the database (i.e. roads, intersections) and address matching are essential to facilitate navigation. This allows queries to be posed regarding routes through the model and ultimately the real world.

Data regarding a road network can be divided into spatial and thematic data. The spatial component of an entity refers to its geometric representation (i.e. point, line, polygon). The thematic component can be viewed as a description of the object and its attributes and serves to give the geometric component an identity (i.e. a highway link with vehicle flow). Geometric data changes much less frequently in modeling road network dynamics, but attribute data may change on very short time scales (Sandell, 1994). Figure 1 shows example data types and their relative temporal update.

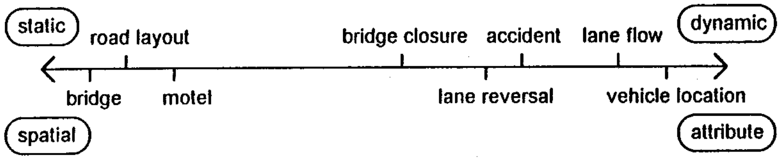

Figure 1. Temporal update scale for example IVHS data.

Distributed Spatial Databases

It it's strictest sense, a distributed database is one in which various parts of a single logical homogeneous database are maintained at geographically separate sites. A distributed database management system is a system that provides transparent access to these data. However, the term distributed database is loosely defined and it is commonly used to refer to any architecture that introduces a degree of interoperability or cooperation between multiple autonomous databases. In general, a distributed database is based on the following primary design concepts:

partitioning - The separating of a database into fragments. Example partitionings in a spatial context are given in Figures 2 and 3. These two partitioning methods, horizontal and vertical, can also be combined into a mixed partitioning.

allocation (distribution) - Allocation involves the optimal distribution of fragments. The question of allocation centers on how to distribute portions of the database among the various component sites. In many cases, the sites exist, and the question is one of integrating the fragments as they are currently allocated.

replication - This is an approach to distributed database design that involves duplicating part (partial replication) or all of a database at various sites to improve access efficiency.

transparency - A distributed DBMS should provide the user with the illusion that all the data is stored locally. This implies that partition boundaries are transparent.

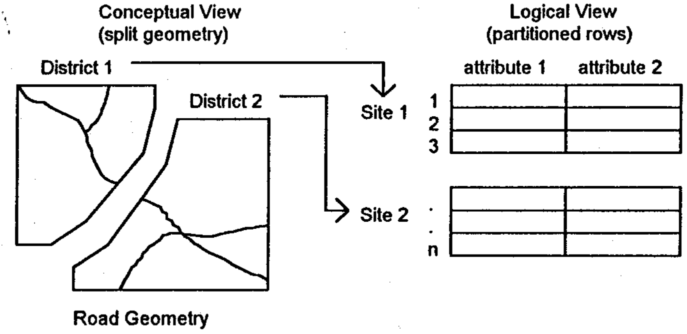

Figure 2. Spatial (horizontal) partitioning.

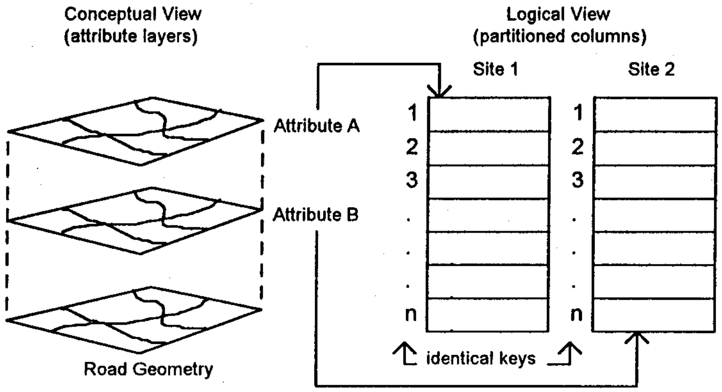

Figure 3. Attribute (vertical) partitioning.

An important issue in the design of a distributed database involves the trade-off between partitioning and replication. As data becomes more dynamic (see Figure 1), there is decreasing motivation to replicate data due to the overhead and complexity of maintaining the replicated copies. Relatively static data, on the other hand, can be replicated at each component site with less difficulty.

REQUIREMENTS

Background

Ultimately, requirements should be driven by users, but who are the users within the IVHS community? Taking the broadest view possible, it would include all parties that have an interest in the location and condition of a region's roads. One classification of users is as follows:

transportation agencies (state, county, and federal DOT)
local governments (cities, counties)
transit authorities (bus, subway, rail, etc.)
private business (delivery services, retailers, trucking companies, etc.)
researchers (academic, private, government)
citizens (commuters, tourists, shoppers, etc.)

All of these parties will inevitably participate in the IVHS database collective and should be considered in the requirements for a community's supporting navigable database. The degree to which a party is responsible for the database determines the influence that the party will have on the design and operation of the database. But what does it mean to participate?

A first step involves defining levels of participation. To date, IVHS projects like Pathfinder and TravTek have focused on introducing a transportation operations center (TOC) in a given region (Summer, 1991) that acts as a vehicular control tower. This approach mirrors the "integrate and centralize" theme that fueled the success of database management systems over the last fifteen years (Litwin, et. al., 1990) and continues to thrive in even the most cutting edge research on very large spatial databases (Guptill & Stonebraker, 1992). These centers serve as collectors and disseminators of information regarding the state of a region's roads. Given that a TOC in a given region provides a community with a centralized IVHS database, there are varying levels of participation, or rights, that can be granted by the TOC:

viewing - Simple viewing and reading of the TOC database (in-car map displays, kiosks, remote log-in, etc.). This is the participation level of in-vehicle navigation systems.

transfer - Transferring TOC data to a local site for local manipulation. This would allow a party to obtain a snapshot of the TOC database for absolute control over the data. This is the dominant sharing model of the spatial data community to date (i.e. "Can I have a copy of that?"). In an IVHS context, this level of participation is not very valuable. Road attribute data is very dynamic, and a database loses its value rapidly if it is transferred to a local site and community-wide updating ceases.

local integration - Integrating any subset of the TOC data with local data. A great number of parties that are interested in the state of a region's roads have their own data to integrate with the TOC data. This participation level holds great potential to serve the needs of parties that have data that is only relevant to them (i.e. academic researchers, private business, local government, etc.). The TOC portion of the database would be read-only, but modification to the local data would be possible.

central integration - Integrating local data into the TOC database for availability to others. Foremost, this includes participants from transportation agencies and the public sector who are primary providers to the TOC regarding a region's roads (i.e. incident locations, directionality, impedances, capacities, vehicle flow, etc.).

modification - Modifying some or all of the TOC database. In the case of a centralized database, access rights would be in the hands of the TOC database administrator. This is a quality control benefit that originally justified establishing the TOC.

From this outline of participation levels, it is clear that the TOC has a high degree of control over the functioning of the database. Participants at the central integration level have to relinquish rights to their data when transferring it to the TOC. To maximize the quality of this data it should remain under local control. It is also doubtful that participation will even exist at the local integration level within the confines of a centralized database. Finally, it is doubtful that a single centralized site can achieve an acceptable level of performance for a statewide or national database. These drawbacks alone motivate examining other avenues for achieving an IVHS database cooperative.

IVHS Database Cooperative (Virtual Road Map)

To temporarily put current technical infeasibilities aside, it is worth defining a vision of the ideal distributed IVHS database cooperative as it may someday exist. This definition can be expressed in terms of requirements, and will serve as a valuable basis for a subsequent evaluation the GIS community in currently meeting these goals. The requirements are as follows:

data sharing - In an IVHS context, data sharing is an absolute necessity. There are data from myriad sources that need to be integrated in real-time, keeping in mind that data quality and value decline rapidly when control moves away the collector. This sharing may be achieved through centralization, replication, or partitioning. Distributed databases strive for efficient data sharing via a divide-and-conquer approach.

local autonomy - It is important that the existing rights of component sites to control their data, systems, and practices are preserved as much as possible. This is in direct opposition to the prior requirement, and a compromise between autonomy and data sharing must be achieved. Historically, most transportation agencies have had complete autonomy, and the direction is towards data-sharing. In striving for this goal, it is important that dependence on a central federated leader, as well as inter-site dependence, be kept to a minimum. This includes the goal of continuous operation, where one site can't crash another site.

integrity - A database intended for vehicle navigation naturally has very high data integrity demands. Partitioning an existing database or integrating data from multiple autonomous sites introduces a host of issues regarding accuracy, integrity, and various classes of heterogeneity (Worboys & Deen, 1991).

efficiency - IVHS data can change rapidly, and the system must also respond rapidly. This arena includes real-time data acquisition and continuous visualization. Data replication can aid in improving efficiency, as some relatively static data road network data may be replicated at each site.

The ideal system to meet these requirements would allow agencies to publish the information that they would like to contribute to the distributed map-database collective, while subscribing to the unlimited amount of information provided by all other sources. Combining all this data under a versatile data model (NCGIA, 1993) would yield a rich virtual IVHS map database. This virtual road map wouldn't physically exist at any single site, but instead, would be the result of a distributed spatial information cooperative or collective. To a user it would simply appear as a single unified map database of the transportation network with a diverse set of themes. A relevant question is then: where is the GIS community in meeting the requirements of such a system?

DISTRIBUTED SPATIAL DATABASE ALTERNATIVES

Three primary GIS architectures appear in the spatial database literature under a variety of names (Vijlbrief & van Oosterom, 1991), but for the purposes of this paper the approaches will be referred to as the dual, shell, and integrated architectures. This naming convention is a variation on Vijlbrief's and van Oosterom's, where "shell" has been substituted for "layer". This substitution was made due to the wide use of the term layer in the GIS community to refer to the thematic partitioning of space. Not all GISs fall cleanly into one of these three categories, but the classification scheme does hold for many systems. Each section provides a description of an approach and its relative potential for distributing a map database.

Dual Architecture (Geo-Relational)

The most common approach to designing a GIS involves separating each geographic entity into its respective spatial and thematic components and storing these two components in separate "dual" databases. Common geographic entities in the context of a navigable database include intersections, road segments, and regions of interest. These two components are linked via a common identifier across the two DBMSs, and retrieving an object involves querying both subsystems through an integrated "umbrella" environment to compose the final answer. This approach also facilitates storing the thematic component entirely within a commercial relational DBMS (i.e. Sybase, Ingres, Oracle). The geographic component, on the other hand, is stored in a proprietary database with its own unique internal access and storage methods. This approach is also commonly referred to as the geo-relational approach, with ARC/Info being the most popular example (Morehouse, 1989).

To date, the most commercially successful vector GISs rely on the dual architecture. These systems have a wealth of experience in serving the needs of the spatial data handling community, and a primary benefit of this approach is its richness in GIS functionality. In this sense, most dual GISs are off-the-shelf systems that allow for the input, management, manipulation, analysis, and output of spatial and thematic information. The history of GIS in transportation applications has developed the capabilities of these systems to a point where many of the required features for meeting the needs of IVHS already exist in some form (i.e. dynamic segmentation, network analysis, dynamic visualization). The question is then: how do these systems measure up in supporting concepts like the transparent sharing of spatial data in real-time?

Figure 4 depicts two different dual GISs at separate sites connected via a distributed, relational DBMS. This example is intended to take a dual architecture as far as possible in meeting the requirements for a distributed, navigable database. The relatively static data (see Figure 1) is stored in the proprietary geometric storage subsystem of each GIS, which precludes the horizontal partitioning of this data (see Figure 2). This drawback may in turn be a plus regarding geometric data integrity, as it requires an entire copy of the geometric portion of the road network to be manually merged and replicated at each site. This approach skirts the integrity problems introduced by spatially fragmenting a road network among autonomous sites (i.e. dynamic boundary matching). The agencies cooperating via this approach would have to join their respective geometry and distribute it periodically via file transfer. The largest benefit of this approach, then, involves the sharing of dynamic attribute data across sites through a RDBMS capable of database distribution (i.e. Sybase, Ingres-Star, Oracle). The RDBMS can be used to vertically partition these attribute tables (Figure 3) across the two example sites, and each GIS can be "fooled" into believing that the attribute data is stored locally. If the GIS at each site has client/server capabilities (von Seggern, 1994), data-sharing can be augmented to include multiple clients at each GIS (server) site. Overall, this approach has the potential to distribute a navigable database, but its limitations must be assessed in a specific prototype implementation.

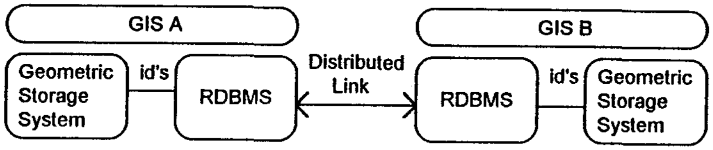

Figure 4. Dual approach - different GISs connected via a distributed RDBMS.

Shell Architecture

Another approach to designing a GIS is the shell architecture. This approach is an attempt to overcome the drawbacks of storing the two components of geographic information in separate subsystems by storing all spatial and thematic data completely within an RDBMS. In this architecture, the spatial data is separated into its basic elements (i.e. points, lines, polygons) and stored in related tables. To retrieve the information, relational joins are performed to reconstruct the required geographic entities. This is accomplished, and made efficient, by placing a spatial support "shell" on top of a standard RDBMS to handle geographic queries. The success of this approach, therefore, hinges on the intelligence of the spatial shell design. An example system that relies on this approach is GEOVIEW (Waugh & Healey, 1988).

This approach may also involve bypassing commercial GISs and developing a special purpose environment directly on top of an RDBMS specifically to meet the requirements of IVHS. This alternative would facilitate tapping all the resources for distributed database management directly, rather than relying on a proprietary commercial GIS shell between the RDBMS and an IVHS application. This circumvents that fact that vendors of shell GISs may not have had time to tap the distributed functionality that has recently emerged in many RDBMS. This inevitably represents a significant amount of programming effort but stands as a viable avenue. Figure 5 depicts two GISs connected via an underlying distributed RDBMS, where the spatial support shell and GIS levels may be commercial or custom.

Theoretically, this approach would allow spatial as well as attribute partitioning of a road network across sites. The GIS could again be "fooled" into believing that the data is entirely stored locally, and the underlying RDBMS would manage the actual physical locations of the partitions. A first step would involve homogenous RDBMSs, but current research into heterogeneous DBMSs may lead to a day when these systems cooperate.

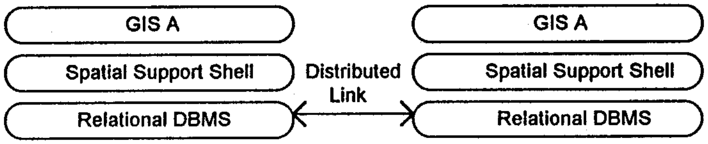

Figure 5. Shell approach - two GISs connected via an underlying distributed RDBMS.

Integrated Architecture

An emerging approach to developing a GIS is the integrated approach. This approach is the least exploited within the commercial GIS community, but it's a popular research focus in both industry and academia. In this approach, the spatial data management and storage facilities are built directly into the DBMS. This precludes the use of a standard RDBMS as the foundation for this architecture, as the facilities for this approach are unsupported by these commercial systems. Instead, an extended-relational or object-oriented DBMS is required. Another driving force behind this approach is the notion that users and application developers can extend the system with their own abstract geographic data types and behaviors. This approach shows the most promise within the research community as the GIS architecture of the future. An example of this architecture is GEO++ (Vijlbrief & van Oosterom, 1991). SMALLWORLD (Newell, 1992) also possesses some of the characteristics of this approach and is a good example of a system that doesn't fall cleanly into one of these three architectures.

These systems represent a sophisticated foundation for the development of a transparently, distributed navigable, database, and support a rich degree of alternatives to implementing a complex geographic database where the integrity of individual geographic entities is preserved. This preservation of integrity is due to the fact that entity components are not divided across two separate databases, but are instead encapsulated together along with their "behavior". These systems are benchmarked much faster than relational databases at managing complex spatial objects (Milne, et. al, 1993), as the spatial data models can be stored in their native form over dividing them or collapsing them into tables. The support for a scalable architecture that allows new servers to seamlessly join a distributed database cooperative with no effect on existing servers is also a major plus in the development of an IVHS database.

The vision behind the integrated approach is an extensible, efficient, spatial data management system capable of supporting concepts expected from a leading-edge DBMS (i.e. transparent database distribution). The object-oriented and extended-relational database markets are fueled by the demand for supporting complex application development in the CAD, CASE, and GIS arenas. IVHS is also targeted as a specific arena for expansion by may of these next-generation DBMS vendors.

At this point, this approach requires a subset of basic GIS development from the ground up, but it is gaining momentum in the IVHS community (JHK, 1993). A system of this type would effectively have to travel the same path as many vector GIS vendors, although not to same degree of generality as the application arena (IVHS) is more specific.

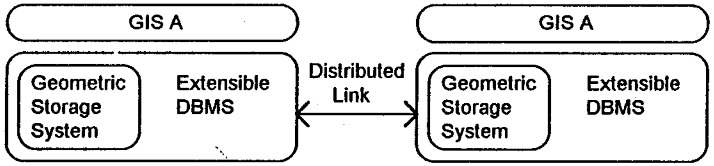

Figure 6. Integrated approach - GIS on a distributed OO/extended-relational DBMS.

CONCLUSIONS

IVHS is raising difficult questions and requirements for the GIS community regarding the sharing of dynamic data. Varying architectural approaches to designing GISs have direct positives and negatives in distributing a spatial database. GISs are relatively mature as single-user planning tools, but there is a wealth of opportunity in linking sites together to achieve spatial database cooperatives. Data collectors are best left as autonomous and separate, but centralization can reduce much of the complexity regarding semantic and system heterogeneity in the interim.

ACKNOWLEDGMENTS

Support for this research was provided by Caltrans through Ramez Gerges at the Caltrans Center for Interoperability (CFI) under grant 65T155. Thanks to Chris Fischer for valuable conversations regarding distributed databases.

REFERENCES

Choy, M., Kwan, M., Leong, H.V. (1993) On real-time distributed geographical database systems. *Proceedings of the 27th Hawaii International Conference on Systems Sciences*. IEEE Computer Society Press, vol. 4, pp. 337-346.

Egenhofer, M.J. (1993) What's special about spatial? Database requirements for vehicle navigation in geographic space. *Proceedings of SIGMOD '93*. Association of Computer Machinery, Washington, DC, pp. 398-402.

Guptill, S.C., Stonebraker, M. (1992) The Sequoia 2000 approach to managing large spatial object databases. *Proceedings of the 5th International Symposium on Spatial Data Handling*, Charleston, SC.

JHK (1993) Database Systems, *TOC Component Evaluation*, Document no. 30, Contract No. 04D922-PC, Emeryville, CA, JHK and Associates.

Lee, N., Karimi, H., Krakiwsky, E. (1989) Road information systems: impact of geographic information systems technology to automatic vehicle navigation and guidance. *Proceedings of the First Vehicle Navigation and Information Systems Conference (VNIS '89).* IEEE, Piscataway, NJ, pp. 337-352.

Litwin, W., Mark, L., Roussopoulos, N. (1990) Interoperability of multiple autonomous databases. *ACM Computing Surveys*, 22:3, pp. 267- 293.

Milne, P., Milton, S., Smith, J. (1993) Geographical object-oriented databases - a case study. *International Journal of Geographical Information Systems*, vol. 7, no. 1, pp. 39-55.

Morehouse, S. (1989). The architecture of ARC/Info. *Auto-Carto 9*, pp. 266-277.

Newell, R. (1992) Practical experiences of using object-orientation to implement a GIS. *GIS/LIS Proceedings*, San Jose, CA, pp. 624-629.

NCGIA (1993) Data Model. *Progress Report - Caltrans Agreement 65T115.* National Center for Geographical Information and Analysis, Santa Barbara, CA.

Sandell, A. (1994) Real-time GIS with continuous, dynamic visualisation of data, design-issues. *Proceedings of EGIS '94.* vol. 2, pp. 1666-1675.

Schiff, T.H. (1993) Data sources and consolidation methods for creating, improving, and maintaining navigation databases. *Proceedings of the Vehicle Navigation and Information Systems Conference (VNIS '93)*, IEEE, Piscataway, NJ, pp. 3-7.

Summer, R. (1991) Data fusion in Pathfinder and TravTek, *Proceedings of the Vehicle Navigation and Information Systems Conference (VNIS '91).* Society of Automotive Engineers, Warrendale, PA, vol. 1, pp. 71-75.

Vijlbrief, T., van Oosterom, P. (1991) The GEO++ system: an extensible GIS. *Proceedings of the 5th International Symposium on Spatial Data Handling*, Charleston, SC, pp. 40-50.

von Seggern, M. (1994) The enterprise GIS: a client/server approach using distributed relational databases to create a multi locational/multi application GIS. *Proceedings of EGIS '94.* vol. 1, pp. 654-660.

Waugh, T.C., Healey, R.G. (1987) The GEOVIEW design. A relational data base approach to geographical data handling. *International Journal of Geographical Information Systems,* vol. 1, no. 2, pp. 101-118.

Worboys, M.F., Deen, S.M. (1991) Semantic Heterogeneity in distributed geographic databases. SIGMOD Record, vol. 20, no. 4, pp. 30-34.