

SPATIAL AND TEMPORAL QUERY SUPPORT FOR GEOSPATIAL DECISION MAKING

Cyrus Shahabi and Craig Knoblock

Integrated Media Systems Center, University of Southern California, Los Angeles, CA 90089

shahabi@usc.edu, knoblock@isi.edu

1. INTRODUCTION

GeoDec is an interdisciplinary project that was started in April 2005 at the University of Southern California Integrated Media Systems Center (IMSC), and consists of a team of computer scientists from the areas of databases, artificial intelligence, computer graphics, computer vision, and systems. The goal of GeoDec is to develop an end-to-end system for the purpose of rapidly and accurately building dynamic information-rich and realistic three-dimensional (3D) geospatial areas (e.g., cities) with a temporal dimension [1]. In addition, by providing interactive spatiotemporal querying, data analysis and visualization capabilities for these geospatial areas, GeoDec aims to enable geospatial decision making for users in a variety of geographic domains, including urban planning, emergency response, military intelligence, simulation & training, computer games, real-estate and news broadcasting.

the correct database layer component and returns results in a format understandable by all our visualization applications.

Because of its emphasis on the temporal aspect of geospatial data, the GeoDec environment can be used by decision makers to visualize object-level *dynamics* as objects change or move. Similarly, the use of web sources where appropriate implies that our environment is automatically updated if the source is changed. Moreover, GeoDec itself is a dynamic system – its modular architecture integrates numerous existing systems, and is designed to scale to include additional components in each layer. The following sections describe the query visualization and analysis features of GeoDec, focusing on Negaah, Jooya, Prometheus and the STDBM, as well as our future work in these areas.

2. GEODEC COMPONENTS

A. QUERY FORMATION AND VISUALIZATION (NEGAAH)

Negaah is a visualization interface for GeoDec that allows the user to navigate and interactively query the 3D environment in real-time, and allows greater decision making flexibility by allowing the user to query geospatial data based on a user-defined selection area. Negaah also allows the user to submit queries based on different time intervals for temporal data such as object trajectories or stored video streams. The user can selectively query and display different layers of information, and move forward or backwards in time. Negaah also supports more sophisticated queries such as calculating the shortest path between two points considering the traffic and other real-time events. Examples of data sources that Negaah can already query or will support in the near future include 3D models, road vector data, raster maps, traffic information, point data (such as building names), moving objects/trajectories (such as trams) and parcel information.

Our future research in this area will focus on allowing the user to form additional types of queries in the user interface. For example, we plan to develop more queries that take advantage of the 3D aspects of the GeoDec environment such as line-of-sight queries. In addition, another research area is providing enhanced data analysis features that reveal correlations, dependencies and other relationships between different geospatial data sources to the user. For example, query results for stored object trajectories (e.g. cars, trams) could also automatically indicate available video streams for the same geographical area and time, thus

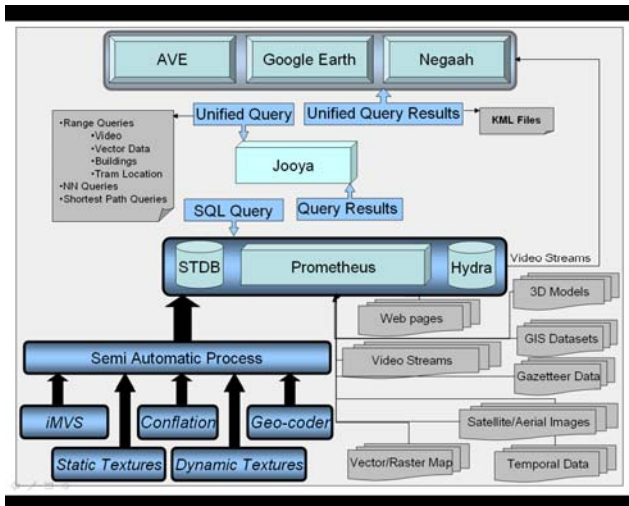


Figure 1: GeoDec System Architecture

Figure 1 depicts the GeoDec system architecture, which broadly consists of three layers. The first is the *visualization layer* that consists of three independent visualization applications namely, AVE [1], Negaah and Google Earth. The *database layer* provides different geospatial data sources via either a physical database such as our spatio-temporal database (STDBM), or a virtual database for web sources such as our information mediator, Prometheus. GeoDec also uses a specialized server (Hydra) for stored video streams. Queries from the visualization layer are directed to a *query middleware layer* (Jooya), which directs the query to

allowing the user to visually connect different data sources. Such queries would be particularly relevant to the intelligence and emergency response domains.

B. SPATIO-TEMPORAL QUERY MIDDLEWARE (JOOYA)

All the queries in Negaah are directed to GeoDec's information mediator/spatio-temporal database component through a middleware layer, Jooya. Jooya offers a universal way of specifying the type of query (e.g., range, shortest path, nearest neighbor, etc.), as well as its parameters and retrieves the results back in a unified format. Depending on the user's query type, Jooya either sends a query to our Spatio-Temporal Database Manager (STDBM), or to Prometheus, our information mediator (both discussed below). Examples of queries sent to STDBM are queries for infrequently updated or bulky data such as road network data, video metadata, and 3D building models. On the other hand, Jooya uses Prometheus to query highly dynamic data such as live traffic information obtained via a variety of web sources.

Jooya returns results in a standard format (specifically, Google Earth's KML format), and this architecture enables any visualization layer that can support KML to sit on top of Jooya for its integrated query and access needs. Jooya, however, would need to include a customized query-interface blade for each new GUI. Hence, Jooya also allows GeoDec to be dynamic at a system-level – new components can be added to the database layer without impacting the visualization modules, and vice versa.

Future research areas for Jooya include performance and reliability improvements through caching as well as developing efficient query plans.

C. INFORMATION MEDIATOR (PROMETHEUS)

GeoDec utilizes a mediator system called Prometheus that provides a uniform query interface to a set of web sources that contain information about a geographic location. These sources range from image or feature vector data sources, such as property parcel data, to more dynamic sources that provide current traffic flows.

There are three main advantages of using the Prometheus mediator system in GeoDec. First, it hides all the different access methods, heterogeneous source schema, and formats. Instead, Jooya can pose queries to Prometheus using a set of global domain concepts, such as roads. For example, there are several data sources that provide road vector data information, such as U.S. Census Bureau's Tigerline files or Navteq. When Jooya directs a query to Prometheus for roads in a given area, Prometheus determines relevant data sources given the user's area of interest and accuracy requirements, retrieves the roads, and returns them to Jooya.

Second, Prometheus can access information from web sources using the Fetch Agent Platform, which allows us to quickly integrate new sources of data, such as traffic information from CalTrans.

Finally, Prometheus also supports database-like queries requiring integration and aggregation of geospatial web sources. A sample query would be to find the sum of the area of all parks within 1000 meters of a parcel. This ability allows GeoDec to dynamically fuse geospatial information in response to user queries.

In future, Prometheus will have the ability to dynamically align different datasets to support more accurate fusion of geospatial sources. This will enable us to quickly integrate new sources into the GeoDec framework without sacrificing the accuracy of the results.

D. SPATIO-TEMPORAL DATABASE MANAGER (STDBM)

As part of GeoDec's back-end, we are also developing a Spatio-Temporal Database Manager (STDBM). This module is in charge of managing spatio-temporal data stored in a database, a task that includes data modeling, storage and retrieval (querying).

Our primary concern in modeling is to design the most efficient model based on the characteristics of the data set as well as the application requirements. For instance, buildings can be modeled as static objects for certain simulations, whereas urban planners may want to model buildings as dynamic objects whose properties (e.g. height) change over time.

For storage, our main focus is now on the storage of bulky vector data such as road networks. Mainly, we want to store vector data at different resolutions with different aggregations so that we can support efficient retrieval given the accuracy needed by a given query.

For retrieval, STDBM provides different features that enable Jooya to retrieve data efficiently from the database. In particular, STDBM supports fast and efficient retrieval of data from huge data sets through the use of spatial indexing, and also provides a set of useful spatial operations. Examples of the latter –which are also supported by the OGS (Open Geospatial Consortium) – include but are not limited to: spatial measurements (e.g. finding the distance between two points), spatial functions (e.g. finding the shortest path), spatial predicates (e.g. true/false queries such as whether there is any tram station located within a mile of a particular building), spatial constructor functions (e.g. create a polygon from set of points) and spatial observer functions (e.g. specific information about a geometry such as the center of a polygon).

3. RELATED WORK

There exist several GIS systems and web-based mapping tools that can access geographic data in

various formats and show different types of geographic data to users by overlaying different data layers.

ESRI software's ArcInfo and ERDAS Imagine are very commonly used GIS systems. The commercial GIS systems allow users to retrieve different types of geospatial data and can handle various formats as well as coordinate systems. The commercial GIS systems can also visualize different data layers and often allow developers to encode various operations, such as, conflation as plug-ins. While the GIS systems are good for retrieving and visualizing data, they provide limited dynamic integration and decision-making capability. The user can retrieve data from various sources and see the data layers on top of each other, but the responsibility of selecting data layers and operations to integrate the layers is on the user. The focus of our work is on dynamically selecting relevant sources, retrieving data from those sources, and integrating the retrieved data using relevant operations.

Google Maps, Google Earth and MSN Virtual Earth are examples of web-based mapping tools that can also display different types of geospatial data on maps and satellite images. Similar to the GIS systems, these systems address the problem of visualizing geospatial data but they do not provide sophisticated querying capabilities to the user. Thus, although we utilize such tools (e.g. Google Earth) to visualize the result of different queries sent to Jooya, we have built our own tools to support querying (i.e. Negaah).

4. CONCLUSION

GeoDec is an end-to-end system to rapidly and accurately integrate heterogeneous geospatial data sources to allow users to visualize and query a geographical region interactively. As discussed in the prior sections, there are several future research areas within each component of this project through which we plan to further enhance the decision making capabilities of the user.

Furthermore, we are also looking at unifying the three different GUIs that currently sit on top of our various databases and middlewares. The first, AVE, is good at fusing video streams with 3D models. The second, Negaah, is targeted towards interaction and querying. And the third one is Google Earth which is a commercial web-based tool. Ideally, we want one GUI combining all three capabilities. However, building such a GUI is challenging. For example, for an efficient video projection, the GUI needs control of the client's resources. This is not possible in a web-based GUI such as Google Earth. The main open problem here is how to efficiently map from terabytes of relevant geospatial and temporal data stored in our various databases (both private and virtual ones) to a limited

number of pixels on the client's display. We are now looking at alternative approaches to do this efficiently.

5. ACKNOWLEDGEMENTS

This research has been funded in part by NSF grants EEC-9529152 (IMSC ERC), IIS-0238560 (PECASE), IIS-0324955 (ITR), and unrestricted cash gifts from Google and Microsoft. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We would also like to thank our following graduate students and research staff helping with this project: Arjun Rihan, Jeff Khoshgozaran, Snehal Thakkar and Ali Khodaei.

6. REFERENCES

1. C. Shahabi, Y-Y. Chiang, K. Chung, K-C. Huang, J. Khoshgozaran-Haghighi, C. Knoblock, S. C. Lee, U. Neumann, R. Nevatia, A. Rihan, S. Thakkar, S. You. GeoDec: Enabling Geospatial Decision Making, IEEE Int'l Conf. on Multimedia & Expo (ICME), Toronto, Canada, July 9-12 2006.

7. ABOUT THE AUTHORS

Dr. Cyrus Shahabi is currently an Associate Professor and the Director of the Information Laboratory (InfoLAB) at the Computer Science Department and also a Research Area Director at the NSF's Integrated Media Systems Center (IMSC) at the University of Southern California. He is also the CTO of Geosemble Technologies. He received his B.S. in Computer Engineering from Sharif University of Technology in 1989 and then his M.S. and Ph.D. degrees in Computer Science from the University of Southern California in May 1993 and August 1996, respectively. He has two books and more than 150 articles, book chapters, and conference papers in the areas of databases, GIS and multimedia. He is currently an associate editor of the IEEE Transactions on Parallel and Distributed Systems (TPDS) and on the editorial board of ACM Computers in Entertainment magazine. He is also a member of the steering committees of IEEE NetDB and ACM GIS. Dr. Shahabi is the recipient of the 2002 National Science Foundation CAREER Award and 2003 Presidential Early Career Awards for Scientists and Engineers (PECASE).

Dr. Craig Knoblock is a Senior Project Leader at the Information Sciences Institute and a Research Professor in Computer Science at the University of Southern California (USC). He is also the Chief Scientist for Geosemble Technologies. He received his Ph.D. in Computer Science from Carnegie Mellon. His current research interests include information integration, automated planning, machine learning, and constraint reasoning and the application of these techniques to geospatial data integration. He has published well over 150 articles, book chapters, and conference papers in information integration, planning, and machine learning. He is a Fellow of the American Association of Artificial Intelligence.