Integrating Geographic Information Systems, Spatial Digital Libraries and Information Spaces for conducting Humanitarian Assistance and Disaster Relief Operations in Urban Environments

Vished Kumar Alejandro Bugacov Murilo Coutinho Robert Neches

University of Southern California, Information Sciences Institute,
4676 Admiralty Way, suite 1001, Marina Del Rey, CA 90292
{vkumar, bugacov, coutinho, rneches}@isi.edu

ABSTRACT

The GeoWorlds system integrates geographic information systems, spatial digital libraries and other information analysis, retrieval and collaboration tools. It supports multiple applications ranging from intelligence gathering to urban planning, to crisis management and response. Teams can rapidly assemble collections of document-based information from the World-Wide Web and other specialized information sources, visualize geo-spatial distribution of these collections and monitor events that might change conclusions or decisions formed on the basis of an initial information set. This functionality is provided within a framework that supports both synchronous and asynchronous collaboration over finding, filtering and organizing information and presenting it in a rich visualization environment.

Keywords

Geographic Information Systems, Digital Libraries, Disaster Relief Operations, Information Spaces, Visualization, Collaboration

1. INTRODUCTION

GeoWorlds is an experimental system that embodies the current results of an ongoing exploration into two closely related issues: expanding the functional capabilities of *geo-spatial information management systems*, and paving the way for *application frameworks* that support rapid instantiation of specialized applications via customization of generic tools.

As an exploration into geo-spatial information management, GeoWorlds demonstrates how carefully integrating three key technologies (the World-Wide Web, Digital Libraries, and Geographic Information Systems), can provide teams of users with a sense of *shared regional vision* -- the ability to marshal and organize everything known about an area, displayed with respect to space and time. It retrieves, organizes and displays available information about a region in rich displays, allowing teams of users in distributed locations to collaboratively assess situations, develop appropriate responses, and monitor the situation's evolution. The system integrates in-house tools, CoTS components, and various products of other research institutes that are collaborating with our effort within the Distributed Collaborative Enterprises Group at the University of Southern California's Information Sciences Institute.

GeoWorlds provides the ability to rapidly assemble a custom repository of information about a geographic area. By selecting a geographic region on a GIS display, GeoWorlds can retrieve, filter and organize sets of documents associated with that region from the Web, and conversely, visualize in the GIS display the geo-spatial distribution of those documents. The system has a notion of a component registry, allowing users to select and initiate registered analysis and visualization tools, in order to present multiple views of a document collection.

All of the above functionality is provided in a framework that enables synchronous and asynchronous collaboration over searching, filtering, and organizing needed information and actions. The collaboration framework supports viewing and annotating the information to facilitate group decision making both to form and record conclusions about the interpretation of shared information, and to select and coordinate actions based upon those conclusions.

Viewed as an application framework exploration, GeoWorlds focuses upon a generic applications class: *situation understanding and management systems.* Examples include systems for disaster relief, urban planning, intelligence analysis, and earth sciences studies. Their common element is the need to help a group develop and maintain their understanding of facts and events pertaining to a physical space over a range of time. We investigate two levels at which frameworks support construction of specific applications: via composition of generic components, and via specialization and organization of information repository contents. This paper reports progress in identifying an appropriate modularization of services and offering service registries that facilitate creating specific applications by mixing-and-matching of service instances. It also reports progress in

ACM GIS '99 11/99 Kansas City, MO USA
© 1999 ACM 1-58113-235-2/99/0011 … $5.00

supporting a collaborative paradigm for creating and populating information repositories, which complements current search and retrieval tools with aids for identifying and organizing the range of topics embedded in an information collection.

In the next section, we give an overview of the GeoWorlds software architecture. In section 3, we describe the system's components and discuss the implementation methodology. Section 4 summarizes our current status, discusses limitations and identifies research issues that have emerged out of GeoWorlds.

2. ARCHITECTURE OVERVIEW

Situation understanding and management applications have a broad range of functional requirements[1].

The requirements that we have identified and considered during the design of the system are:

- Tools for forming and analyzing collections of information, structured for dynamic storage, retrieval, processing and analysis of data and information available about the region/situation.
- Simulation of spatial/temporal processes relevant to the situation (e.g. weather, transportation, logistics, etc.).
- Visualization and analysis of the information and results.
- Marking, annotating and collaboration of the information between spatially distributed teams.

A major challenge is to provide a framework into which the best available tools can be integrated with minimal effort and in which these components can be connected to each other to create new applications. With this in mind, we designed GeoWorlds as a component-based system that will be able to support a continuous increase of functionality, as new and more sophisticated tools become available.

As shown in Figure 1, on one level the GeoWorlds system is implemented using a client-server architecture paradigm. On another level, it is a distributed system architecture where registered services such as Data Warehouses and Analysis tools can be running on multiple widely distributed hosts.

GeoWorlds defines a set of *interfaces*, contracts that external services must adhere to enable registration with the system. Also included are interfaces for handling core GeoWorlds' data structures like the document model, which defines a model for data exchange between the analysis and visual components of the system.

2.1 GeoWorlds Client Server Overview

The server is composed of five main modules or components. The *GeoWorlds Kernel* comprises a core set of functionality required for communication and data transfer between the other components. The *Collaborative Server* module handles synchronous collaboration between the collaboration-enabled

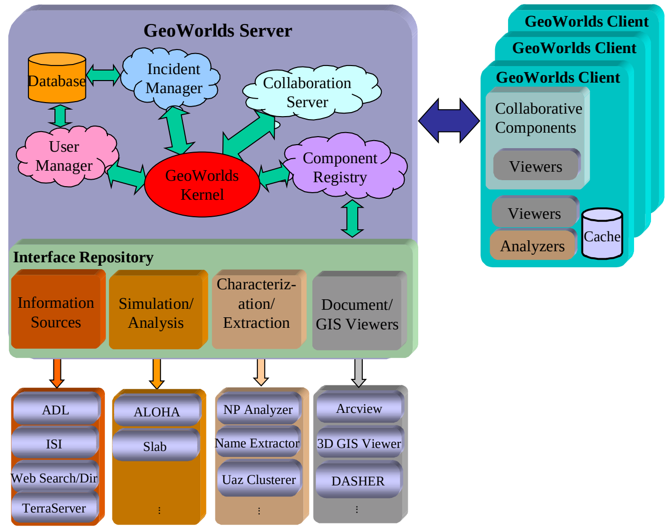

Figure 1: The GeoWorlds System Architecture.

components of the GeoWorlds Clients. There can exist multiple servers servicing the needs of multiple sets of collaborative clients. The *Incident Manager* maintains a database of both current and past incidents, providing hooks for the other parts of GeoWorlds to access this information. The *User Manager* handles user profiles, authentication and security for the GeoWorlds system. The Component Registry allows application developers to make new services and components available, as described in Section 2.2 below.

The client side is embedded in a collaborative framework, and therefore contains applications and services that support collaboration. The client provides a user interface to log in to the server and start collaborative sessions to share information spaces with other clients in the system. It could also contain client-specific non-collaborative services installed locally on the system.

2.2 Component Registration

A valuable concept in component-based distributed systems is the possibility of integrating new components and registering with the system the services that they provide. This allows the users and/or the other components to become aware of the requirements and resources that this new component brings into the system.

In the current version of GeoWorlds, services are registered in the GeoWorlds server and their information is available to all clients joining a GeoWorlds session. The Component Registry provides access to components through the Interface Repository, which specifies the known classes of components and their interfaces. A new component is registered by specifying the class of components it belongs to (or defining a new class and its interface), and then registering a particular implementation of the interface associated with that class.

---

[1] HA/DR (Humanitarian Assistance and Disaster Relief) operations belong to this set of applications and the current version of GeoWorlds is intentionally biased in that direction [8].

3. IMPLEMENTATION ISSUES

The GeoWorlds system integrates technologies from several universities and research labs, each of them contributing with the reference implementation of one or more components from the basic set of components the architecture describes.

The core functionality of GeoWorlds is implemented in Java 2. Java wrapper classes are implemented to encapsulate non-Java components and COTS software (such as ESRI's ArcView). Communications between components was achieved using Java RMI, TCP/IP sockets or Windows dynamic data exchange (DDE), depending on the component being integrated with the system. The following sections describe the set of reference implementations of the components integrated in the system so far.

3.1 Incident Registration

In certain applications, the registration of incidents or events may play an important role in providing a context within which teams work. Incidents are characterized by their type and their location, which is entered as a street address (although it need not be complete). Once the incident has been defined, the system is able to map its location in a GIS Viewer. In case the location has been described by an address, GeoWorlds looks up registered services implementing a *GeoCodeServer* interface and queries each of them in turn until appropriate Latitude-Longitude coordinate locations are returned. Geocoding is the process by which street addresses are mapped to geographic coordinates.

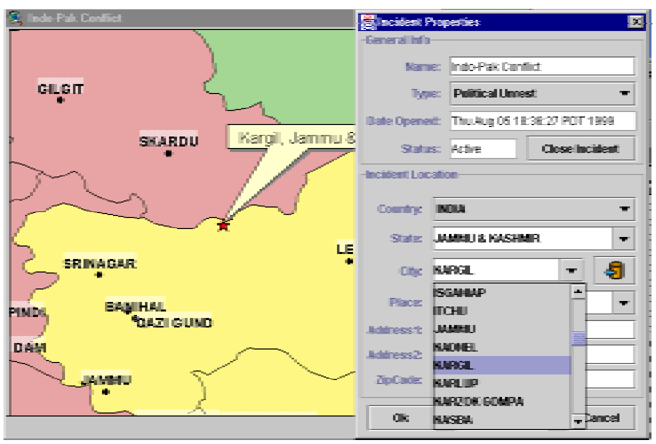

Figure 2: Incident Registration.

3.2 Geographic Information System

Instantiating multiple applications with GIS raises two issues: 1) Viewing and presentation of traditional GIS information and 2) Extensions beyond conventional GIS.

Viewing and presentation of traditional GIS information. The current version of GeoWorlds has two GIS Viewing and Analysis tools. Conventional visualization of the geographic information and other typical GIS functions can be performed using ESRI's Arcview software. We have also initiated work towards implementing a 3D GIS Viewer that would allow the navigation of a geographic region and its terrain, adding another dimension towards visualization of the geo-spatial distribution of document collections.

Arcview is registered as a GIS component that has been integrated into the system using a Java wrapper implementing the GeoWorlds GISViewer interface and acts as the receiver for all requests to Arcview from GeoWorlds. The implementation consists of DDE (Dynamic Data Exchange) calls to invoke scripts, written in Avenue (Arcview's own scripting Language), via a Java-DDE bridge. These scripts are bundled in an Arcview project that is loaded in when the GeoWorlds Server launches Arcview for the first time. This implementation allows the system to automatically load into Arcview all the layers of geographic information that has been downloaded from the data repositories.

The 3D GIS Viewer component is implemented in Java and processes digital elevation maps in DEM format overlayed with aerial orthophotos, additional layers of geographic information (e.g., ESRI's shape 3D format files) and clusters of spatially-referenced documents to generate a 3D scene in either Java3D or VRML. In Figure 6 we show a VRML 3D scene with terrain and satellite information of the San Diego Harbor region where the white glyphs represent the geographic location referenced by a collection of documents that were clustered using our document analysis component. In addition, these forms have sensors that can be used either to signal the geographic location associated with the selected glyphs (red light shown in Figure 7) or launch some external application (e.g., a Web browser) to inspect its associated information.

In the current state of the system, the 3D-viewer processes static terrain and geographic data for a particular region but we are expanding our capabilities to enable the system to dynamically generate a 3D representation of a selected region. Our initial effort in the 3D GIS Viewer generated the 3D scene in VRML but it has been extended to Java3D providing better integration between the viewer and the rest of the components.

Extensions beyond conventional GIS. Support for the interaction between traditional GIS components and document search and retrieval is a key part of GeoWorlds. It allows the user to interactively move back and forth between document-oriented and geography-oriented views of the situation and interact with other system components such as the toxic-gas diffusion analysis tool and the document analysis tool DASHER (see Section 3.4). For example, using ArcView the user can select some layers and extract the list of place names (e.g., streets, hospitals, airports) whose location fall inside a marked region. This region can either be selected by the user or it may be the region of space which is being affected by some event occurring in the region, for example the result form the model prediction of the spread of a toxic material. The list of place names is then exported to DASHER where the user can search DASHER's information spaces to obtain documents associated to the extracted place names. Flow of information in the opposite direction is also possible. The location information in the form of street addresses can be extracted from documents originating from structured sources, such as online Yellow pages. These addresses are then geocoded and sent to the GIS display for geo-referencing the selected documents. The system currently supports location extraction from documents that are semi-structured, that is, from documents that have parts pre-assigned for address location (such as GTE yellow page documents.) A more general location extraction scheme is also supported, using natural language extraction and summarization tools for parsing documents and obtaining names of geographic entities the documents refer to. The coordinates of these entities are then obtained from a query to a GeoNameServer service registered with GeoWorlds. Our current implementation of the Named Entity Extractor recognizes direct references to place names, but more sophisticated extraction tools which recognize indirect geographic references could be implemented and plugged into GeoWorlds just as easily.

3.3 Analysis Component

To coordinate operations, users of GeoWorlds must be able to obtain predictions of the physical processes involved in the situation. The system needs direct interaction with analytical and simulation tools that can properly model these processes in space and time and compute predictions.

So far, we have integrated into GeoWorlds, ALOHA, a toxic-gas plume analysis tool [1] for modeling the dispersion of heavier-than-air gases. ALOHA takes into account the atmospheric conditions and properties of the terrain to compute the time-averaged concentration of the toxic substance as a function of time and downwind distance. It then calculates the maximum possible extent of the region that will be effected by the plume. This extent can then be overlayed over other layers in the GIS Viewer to aid in visualization of the effected areas.

3.4 Document Management Component

GeoWorlds uses the USC/ISI DASHER Information Space Analysis Tools [6] as its document management component. DASHER (shown in Figure 5) helps users rapidly gather information on a given topic by providing tools to find, filter and categorize documents returned as results of queries to various information sources. These categorized collections of documents can be stored as *Information Spaces* in customized repositories, to be made available during emergencies.

DASHER implements wrappers to a large number of search engines, meta-search engines, yellow pages and other web-directory services such as Yahoo! giving users access to a large percentage of information out on the web.

Within its framework, DASHER also provides access to a large number of analysis tools, including University of Arizona's Noun Phrase analysis tools [3], providing the users with a better understanding of their data.

3.5 Collaboration Components

GeoWorlds features synchronous collaboration and joint information exploration capabilities by incorporating University of Illinois NCSA's Habanero [2] as a component. Habanero provides the mechanisms by which teams of spatially distributed users of GeoWorlds can see in real-time what each of the remote users are doing in terms of exploring and analyzing the different information spaces that they are working on.

The GeoWorlds client is embedded in the Habanero framework to allow GeoWorlds client sessions to collaborate by exchanging mouse and keyboard events. Typically, the Habanero server is run on the same machine running the GeoWorlds server, but this condition is not restrictive. Habanero clients can define a new collaborative session or join an existing one.

Though we have implemented collaboration in GeoWorlds via Habanero, we are also looking at other collaboration frameworks. Collaborative frameworks tend to be optimized for certain applications and our vision is to be able to replace one framework with another, depending on the task. On the other hand, each framework has its own unique way of adding collaboration to components thus building a dependency on the framework, making it difficult to have a generalized component-based solution to this problem.

3.6 Structured Information Repositories

GeoWorlds' provides access to large structured information sources such as Digital Libraries and/or real-time sensor data from spatially distributed sources by exposing a *DataWarehouse* interface as part of its API. Data sources that need to integrate with GeoWorlds register wrappers that implement this interface. These data sources can be present both locally or distributed on a network and all communication with remote repositories is achieved through Java's Remote Method Invocation. In the current version we have created wrappers for UC Santa Barbara's Alexandria Digital Library [5] and USC/ISI's own geo-spatial Data Warehouse. Thus the users of GeoWorlds can transparently query multiple data sources and get back results independent of their structure and format. This feature extends the amount and type of geographic data and information that GeoWorlds users have access to as opposed to typical GIS systems that work with local (static) non-distributed data sources.

GeoWorlds' default implementation of the *DataWarehouse* interface comprises an Oracle Relational Database Server and a Web Server for serving maps to the client. The Oracle Database has been structured to store tables containing meta-data on the maps served and gazetteers, which acts as GeoWorlds' Name Querying Service. On registering an incident with GeoWorlds the user queries the Data Warehouse and retrieves meta-data to help select geographic data for download and subsequent viewing in one of GeoWorlds' registered GIS Viewer's.

GeoWorlds provides a simple query tool via which queries can be defined in two ways - by specifying place names or by specifying a geographic extent (see Figure 3). The results of these queries are presented to the user in a "Thumbnail Browser" as shown in Figure 4. This browser displays thumbnails (if available) of the maps/documents that satisfied the query. Also displayed is available meta-data associated with the document and made available by the information source. It is quite unlikely, that different digital data sources agree on a similar structure for their meta-data. GeoWorlds defines a *DataWarehouseDocument* interface, which is the data model every registered Information source must return its data in. Through this mechanism the user can be presented with data in a unified format regardless of the data source. GTE BBN's OpenMap GIS library implements the Map Query Component.

4. CONCLUSIONS AND FUTURE WORK

The GeoWorlds system, as it stands today, is best understood as an initial experimental effort to demonstrate how the ability to coordinate geographic information with document collections, and to collaboratively view and discuss the related information, can greatly enhance the situation understanding process. The current version of the system aims at helping teams performing humanitarian assistant/disaster relief operation to better understand how their situations play out in space and time. These capabilities are embedded in a framework that enables synchronous and asynchronous collaboration between teams of users.

GeoWorlds has been installed for experimental use at US Pacific Command's Virtual Information Center in Hawaii. Based on feedback, we have added a number of features that enable GeoWorlds to integrate extremely well with their information gathering and processing pipeline. Other enhancements requested and candidates for future work are the integration of live audio/video capabilities to the system, ability to work with different languages and access to more information sources.

Our goal for GeoWorlds is to have a system where new available components can be added in a plug-and-play fashion and those with similar functionality can be swapped with one another at run-time. We are looking at component-based distributed architectures like CORBA and JINI as candidates for implementing GeoWorlds' Component Registry. More work needs to be done on enhancing the interaction between document views and geographic views and giving the user an immersive 3D environment containing digital terrain information, aerial or satellite images as textures, results from numerical simulations and mapping of the document space.

GeoWorlds has raised a number of interesting and difficult issues in Software System Architecture, Geographic Information Systems, Natural Language processing and Digital Library Technology. How do different components in the system discover services and data resources available to them? How can geo-spatial objects from different sources and projections interact with one another in a common view or context? How can we extract from documents indirect or context dependent references to geographic entities? How can information sources provide to a system a measure of the quality and quantity of data they possess on a certain topic? These are only a few of the many questions that remain unanswered today, and are topics of active research.

5. ACKNOWLEDGEMENTS

A system like GeoWorlds wouldn't be feasible if it wasn't for the help and collaboration of many people. We would like to thank all research teams involved in this effort. In particular, we would like to thank Mike Sullivan (BBN/GTE), Hsinchun Chen, Dorbin Ng (UAz), Robert Wilensky, ByungHoon Kang (UCB), Larry Jackson, Terry McLaren, Al Woss (NCSA), Terence Smith, Jim Frew, Greg Janee (UCSB), Bill Pottenger, Nancy Yeager (UIUC), John Wilson and Ed DeYoung (USC), Ke-Thia Yao, Ragy Eleish, In-Young Ko (USC/ISI). We would also like to thank Ron Larsen for his support and encouragement. This work is sponsored by DARPA under Fort Huachuca Contract No. DABT63-95-C-0101.

6. REFERENCES

[1] Aloha (Areal Locations of Hazardous Atmospheres): Chemical Air Dispersion Model. Distributed and supported by the National Safety Council (www.nsc.org/ehc/cam/alohapg.htm).

[2] Chabert, A., Grossman, E., Jackson, L., S., Pietrowiz, S., R., and Seguin, C., "Java object-sharing in Habanero," Commun. ACM 41, 6 (Jun. 1998), Pages 69-76

[3] Chen, H., Schuffels, C., and Orwig, R., "Internet Categorization and Search: A Machine Learning Approach," Journal of Visual Communication and Image Representation, Special Issue on Digital Libraries, Volume 7, Number 1, Pages 88-102, 1996.

[4] Frew, J., Freeston, M., Kemp, R., Simpson, J., & Smith, T. (1996). The Alexandria Digital Library testbed. D-Lib Magazine (July/August 1996). www.dlib.org/dlib/july96/alexandria/07frew.html

[5] Neches, R., Abhinkar, S., Hu, F., Eleish, R., Ko, I., Yao, K., Zhu, Q., Will, P., Collaborative Information Space Analysis Tools, D-Lib Magazine, October 1998. For additional information see www.isi.edu/dasher.

[6] Phelps, T., A., and Wilensky, R., "Toward Active, Extensible, Networked Documents: Multivalent Architecture and Applications", Proceedings of Digital Libraries '96.

[7] M. Coutinho, R. Neches, A. Bugacov, V. Kumar, "GeoWorlds: A Geographically-based Information System for Situation Understanding and Management", to appear in GeoInformatica, Proceedings of TeleGeo '99, Lyon, France, May 6-7, 1999.

[8] Ermak, D., L., User's Manual for Slab: An Atmospheric dispersion model for denser-than-air releases. June 1990. Lawrence Livermore National Laboratory and Lakes Enviromental (www.lakes-enviromental.com).

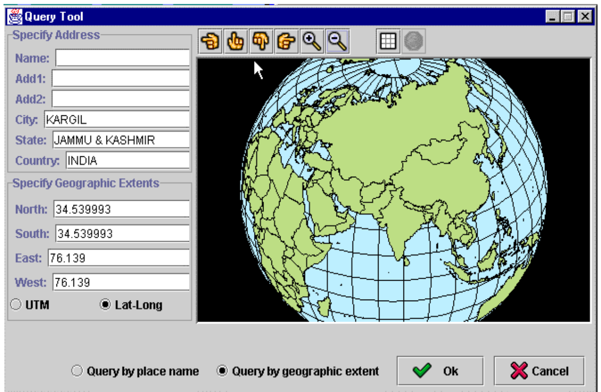

Figure 3: A snapshot of the GeoWorlds Query Tool.

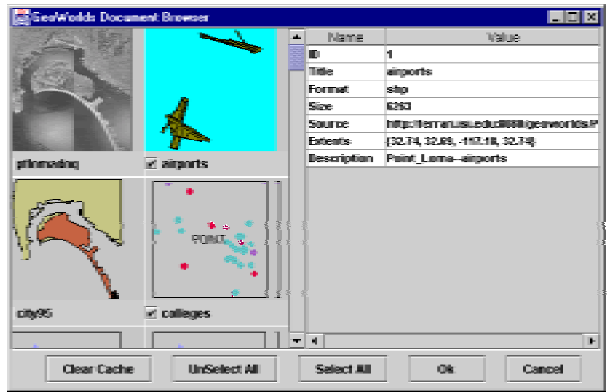

Figure 4: The Thumbnail Browser for presenting geographic query search results to the user.

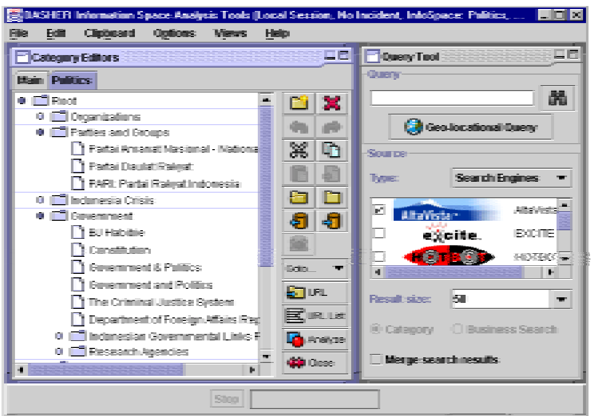

Figure 5: The DASHER User Interface showing the Query Tool on the right and a category tree view of a document collection.

Figure 6: The 3D GIS Viewer showing the region represented by the document cluster highlighted in red and the corresponding documents in the DASHER window.

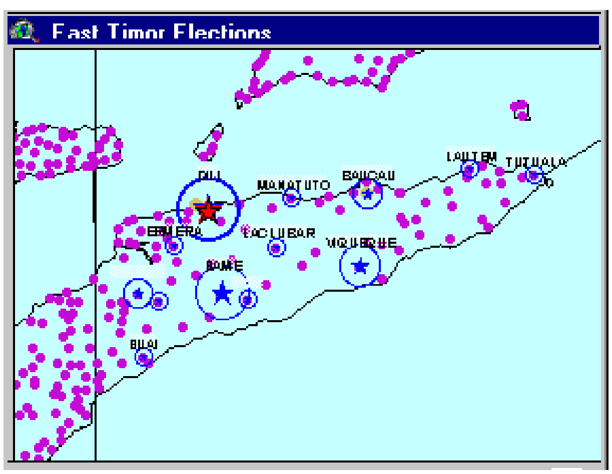

Figure 7: The GIS Viewer showing a document collection and geo-located on a map (indicated by the blue circles). The sizes of the blue circles are proportional to the number of documents found referencing that region.

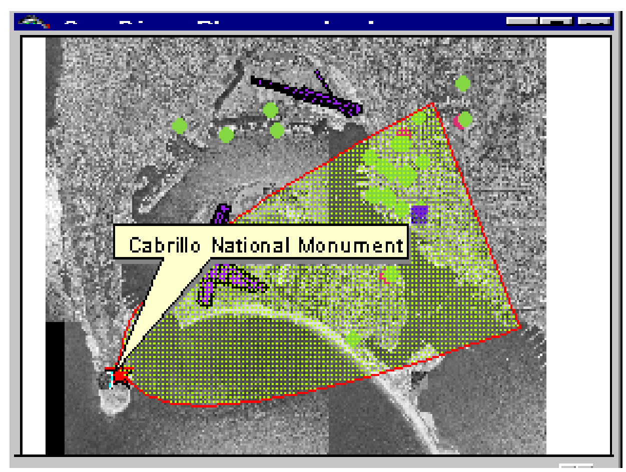

Figure 8: A snapshot of the GIS system showing the results of a Phosgene Leak Analysis overlayed on top existing map layers.