InfoStar: An Adaptive Visual Analytics Platform for Mobile Devices

Antonio Sanfilippo[1], Richard May[1], Gary Danielson[1], Bob Baddeley[1], Rick Riensche[1], Skip Kirby[2], Sharon Collins[2], Susan Thornton[2], Kenneth Washington[3], Matt Schrager[3], Jamie Van Randwyk[3,] Bob Borchers[4], Doug Gatchell[5]

[1] Pacific Northwest National Laboratory
PO Box 999, Richland, WA 99352, USA
{Antonio.Sanfilippo, Richard.May, Gary.Danielson,
Robert.Baddeley, Rick.Riensche}@pnl.gov
[2] East Carolina University
Greenville, NC 27858, USA
{KIRBYC, COLLINSS, Thorntons}@mail.ecu.edu
3 Sandia National Laboratories
P.O. Box 969, Livermore, CA 94551, USA
{kewashi,maschra,jvanran}@sandia.gov
[4] Maui High Performance Computing Center,
550 Lipoa Parkway, Kihei, Maui, Hawai'i 96753, USA
bob.borchers@mhpcc.hpc.mil
[5] National Science Foundation
4201 Wilson Boulevard, Arlington, VA 22230, USA
dgatchell@nsf.gov

Abstract. We present the design and implementation of *InfoStar*, an adaptive visual analytics platform for mobile devices such as PDAs, laptops, Tablet PCs and mobile phones. *InfoStar* extends the reach of visual analytics technology beyond the traditional desktop paradigm to provide ubiquitous access to interactive visualizations of information spaces. These visualizations are critical in addressing the knowledge needs of human agents operating in the field, in areas as diverse as business, homeland security, law enforcement, protective services, emergency medical services and scientific discovery. We describe an initial real world deployment of this technology, in which the *InfoStar* platform has been used to offer mobile access to scheduling and venue information to conference attendees at Supercomputing 2004.

Introduction

We value technology-enabled mobility because it increases our productivity, it enhances our ability to express ourselves, it gives us a sense of safety and security, it enables ubiquitous exploration and discovery, and it supports our aspirations for communal belonging (Stoop, 2003). In this paper, we demonstrate how the combination of pervasive wireless communication and visual analytics provides a technology-

enabled mobility framework which can be made to satisfy the knowledge needs of human agents operating in the field.

Currently, visual analytics software tools are limited to stationary knowledge discovery tasks because have been designed for desktop systems and have therefore not been used to address the knowledge needs of professionals operating in the field such as first responders. The goal of this paper is to begin exploring how such a limitation can be removed by presenting the design and implementation of an adaptable visual analytics platform that extends the reach of visual analytics technology beyond the traditional desktop paradigm to provide ubiquitous access to interactive visualizations of information spaces on mobile devices.

Visual analytics

Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces. It is a growing area in the information management and knowledge discovery arena, where interactive visualization is used to represent and manipulate relations between concepts in data sets. The IN-SPIRE™ [Hetzler and Turner 2004] and Starlight [Rex et al. 1999, Risch et al. 1997] tools developed at Pacific Northwest National Laboratory (PNNL) provide an example of visual analytic platforms.

IN-SPIRE™ (http://in-spire.pnl.gov) can quickly and automatically convey the gist of large sets of unstructured text documents by clustering similar documents together to unveil common themes and reveal hidden relationships within the collection.

Starlight (http://starlight.pnl.gov) information graphics are classified according to the type of information they portray. The two major divisions are *non-spatial information graphics* such as text and numeric data, and *inherently spatial information graphics*, which depict information containing intrinsic spatial coordinates, such as geospatial or CAD data. One of Starlight's unique capabilities is its ability to integrate these two types of information so that they may be analyzed together.
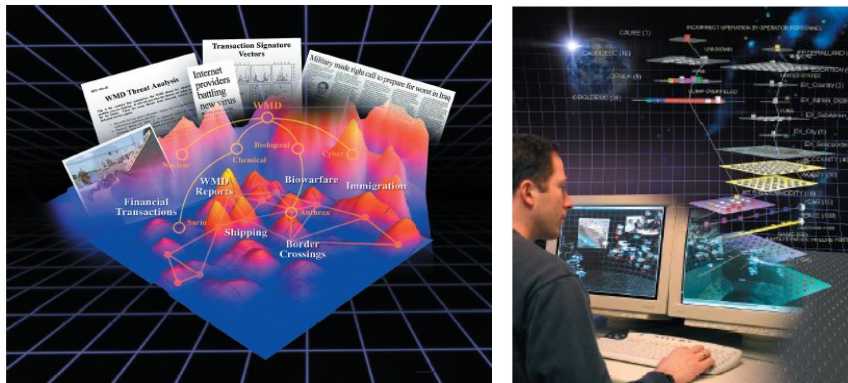

Figure 1: Displays from the IN-SPIRE™ (left) and Starlight (right) tools developed at PNNL.

Visual analytics on Mobile Computing Platforms: Needs and Challenges

The creation of visual analytics for mobile environments has several challenges. Some of these challenges are common to general research in ubiquitous computing. For example, the support of interactive graphics with region picking, hyperlinks and mouse-over events (Tarasewich 2003) is an important issue for the design of interfaces for ubiquitous computing devices. Connectivity and bandwidth is also a critical challenge as visual analytic applications can involve large (up to hundreds of terabytes) data sets and require extensive computational needs. Finally, the issues of screen size, resolution, and color depth as well as different and sometimes limiting input devices pose challenges just as they do for any complex interactive application (Jones 2000, Buchanan 2001, Lee 2003, and Ericsson 2001).

Some of the challenges more specific to visual analytics include the following.

- <u>User Information needs</u>: There is a wide variety of information needs dependent on the type of user and the task to be performed. This requires a careful scoping of the problem space being addressed and a realization of whether and how the results apply to other situations (Campbell 2004).

- <u>Multiple Situational Environments</u>: Users do not have the luxury of using applications only in controlled environments such as a parked car. Interfaces need to support a range of techniques to accomplish the same task in less favorable contexts (e.g. while walking, in dark environments). Selection of the appropriate interaction technique depends on the needed level of control and situational capability of the user to use the interface (Campbell 2004).

- <u>Adapting Existing Techniques</u>: Established visual analytics techniques have been developed for desktop systems. Adapting these for use on ubiquitous devices represents a potentially large set of unexplored challenges.

InfoStar at SC2004

To provide a concrete example of the technology under analysis, we will focus on an initial real world deployment in which the *InfoStar* platform has been used to offer mobile access to scheduling and venue information to conference attendees at Supercomputing 2004 (SC2004).

Requirements

The development of *InfoStar* for SC2004 was guided by requirements that presented various challenges, due to the novelty of the task and in particular the integration of interactive visualization on devices with limited screen real estate. Here is a list of the major requirements we addressed.

- <u>Casual Devices and Users</u>: We did not have prior access to install software on mobile devices nor to train users on the application. The application had to

adapt to the device of the user—be it a desktop, laptop, Tablet PC, pen-based PDA, jog dial-based PDA or cell phone—without manual intervention. A single version of the application should run on all devices and "degrade gracefully" on devices that have limited capabilities. Consequently, standards and simple layouts should be used to allow for maximum compatibility across all devices.

- Discourse Appliances: The application should not place any taxing storage, processing or networking requirements on the target platform, e.g.: all conference content should be sitting on a central database, any content processing should be done on a server, and download size should be as small as possible to accommodate devices that charge per KB of data transferred.
- Interface Design: The interface should provide layers of dynamic details with intuitive navigation and effective "drill down" capabilities providing multiple ways of accessing information. Incorporating existing visual exploration should be a key facilitator for rapid access to information.

System Architecture

The next step was to choose the technologies that would best fulfill these requirements. Since most devices are capable of retrieving URLs and have a built in web browser, HTML was the most appropriate option to display text. We decided to use CSS for formatting and layout since devices that could not render graphics or formatting would simply ignore the CSS and the HTML would remain intact.

For the graphics part of the application, many options were considered, including Macromedia Flash, SVG, VRML, and dynamically generated JPEGs. Prima facie, SVG and VRML seemed to be strong candidates. However, on second inspection we realized that implementations of SVG and VRML are currently not available for most mobile devices. We hope that this situation will soon change, but for our immediate needs we had to look at other options. We entertained the possibility of generating images dynamically, but quickly rejected this idea as it would have resulted in limited interactions and more stress on the web server. We finally chose Macromedia Flash as it offers graphic support across multiple platforms, and it is generally already installed in most devices, including some PDAs. Macromedia Flash provided most of the functionality we were looking for: the ability to query web services, manipulate XML results, display complex graphics, and perform complex interactions.[1]

The *InfoStar* platform interfaces to a wireless network and includes four distinct component layers: data feeds, data aggregation, data analysis and data presentation.

---

[1] We did not pursue graphic support on Blackberries or cell phones since we did not find any graphics solution that would display complex dynamic graphic data on these devices. Moreover, the lack of pen-based input on these devices would have made complex interactions difficult. We anticipate that as the *InfoStar* platforms evolve, we will be able to provide graphic support for these devices as well. For the time being, our focus has been on a degrading gracefully to text-only representations for these devices.

Networking

Each year, a high-performance network named SCinet is built at the Supercomputing conference to provide connectivity to conference attendees, vendors, exhibitors and participants in technical challenges. The SC2004 SCinet network took advantage of the pervasive fiber infrastructure available at Pittsburgh's David L. Lawrence Convention Center to provide high-speed wired and wireless access throughout the building. The SCinet wireless network supported up to 1,300 simultaneous users through the 802.11a/b/g protocols.

The *InfoStar* server was connected at 100 Mbps via the SCinet wireless network and had no problems serving data to conference attendees. Additionally, the IP address space used for the conference was routable Internet address space, making the *InfoStar* server publicly available. Conference attendees were therefore able to access *InfoStar* from remote locations with an Internet (wired or wireless) connection such as their hotel room or home locations. Figure 2 provides a diagrammatic representation of the interface between the *InfoStar* server and the SCinet wired and wireless network.
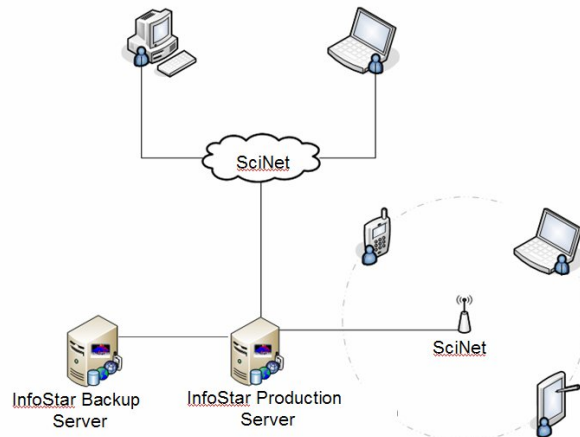

Figure 2: The *InfoStar*-SCinet interface.

System Components

A graphical summary of the *InfoStar* system components is given in Figure 3. The first component layer is the data feed layer. Data feeds for SC2004 came from a variety of sources, including XML streams and spreadsheet files. All data feeds were stored in a Microsoft Sharepoint database which provided a one-stop content access point for all *InfoStar* components and conference personnel who needed special data access (system administrators, conference executives and content owners). Data feeds ranged

from static files updated periodically to interactive databases where exhibitors could enter data regarding events in a "just in time" fashion.

The second layer is the data aggregation layer. Data from the various feeds are collected by a custom application (the "Ingest Manager & Data Loader", see Figure 3 below), and combined into a single source XML file. The data is retrieved using web services. This layer involves renaming/modification of key attributes to ensure a consistent format for all events.

The third layer is the data fusion and analysis. At this layer, the Ingest Manager application feeds the aggregate XML file into PNNL's IN-SPIRE software for contextual analysis. IN-SPIRE groups abstracts into clusters based on its analysis of their topical similarity, and it returns a set of two-dimensional coordinates for graphical representation. The Ingest Manager stores these coordinates along with event data such as times and locations. A table then maps the event time and location with the clustered abstract information in a SQL server database (in tandem with a raw XML file).

The fourth layer is the data presentation layer. Data is presented to users in a variety of forms including text-only representations, or for more capable client platforms, a Flash-based "Visual Exploration" page. As in the data aggregation layer, the presentation layer receives its data from the Ingest Manager application by way of web services. The use of web services for data transfer between the various layers allows for maximum flexibility. For example, we ran the four component layers on a single machine, but were ready to move the data aggregation layer to a second machine should the computational load prove too heavy for a single machine.
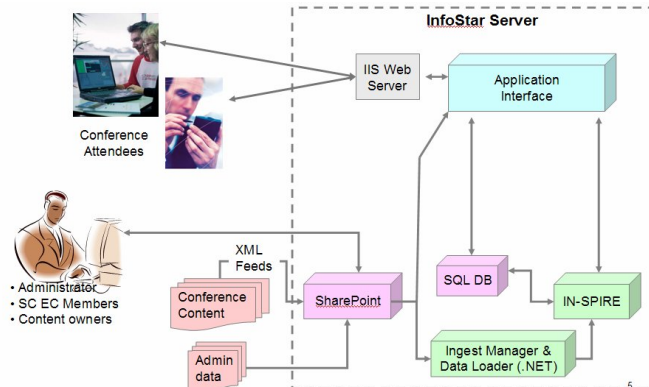

Figure 3: *InfoStar* system components.

Features and Functionality

The SC2004 *InfoStar* system features an interface with a streamlined layout comprising five major sections and several minor sections divided by task or feature, as shown

in Figure 4. The web pages and the attached cascading style sheets are valid W3C documents.
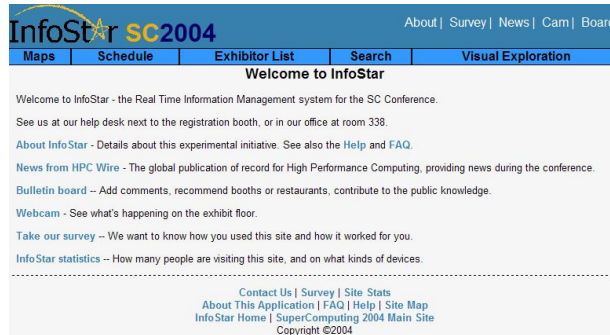

Figure 4: Entry page in the SC2004 *InfoStar* interface.

The major sections (Maps, Schedule, Exhibitor List, Search and Visual Exploration) provide most of the information about the conference. Other sections offer attractive additions or standard features such as "help", "about", contact pages and a real time view on the exhibition floor (via web cam).

The function of the first major section is to display maps of interest to conference attendees, e.g. the layout of the conference halls and the exhibitors' area. Maps were provided for all major areas in the conference center and each room on a map had links to information about events occurring in that room. This included the technical venue and the exhibitor show floor. Figure 5 shows an example map with a room highlighted with a star.
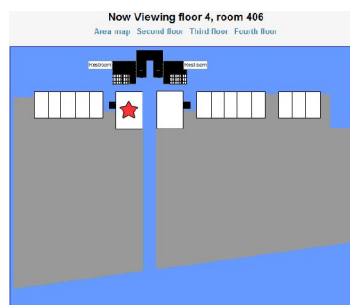

Figure 5: Finding a given room in the Map through a location link.

The second major section provides scheduling information. Users could select conference events based on the day or events currently underway. The initial list showed summary information but provided access to all information on the event.

The third section provides information about conference exhibitors. Exhibitor content, including sponsored events, was accessed through an alphabetical list. Exhibitors provided and maintained their own information

The fourth section allows users to search events by type and time, including events in progress.

The fifth major section was the visual exploration page which incorporated an existing visual analytic technique. This provides an interactive scatter plot view of the events taking place at the conference. Events that are more strongly related are closer together on the plot. Groups of strongly related events are referred to as clusters. Using Flash and the output from the IN-SPIRE engine, events can be displayed as a 'galaxy' of elements that are color-coded by event type. The user can zoom in to view clusters of documents in greater details and restrict the scope of the information displayed by selecting specific days and event type, as shown in Figure 6.
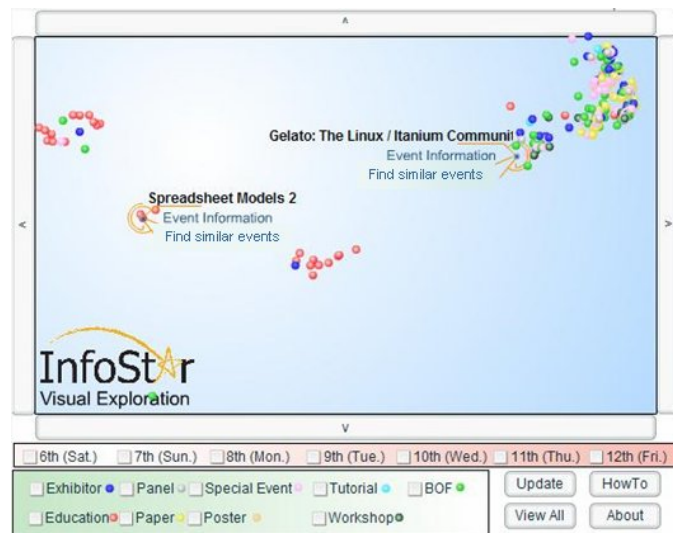

Figure 6: *InfoStar* cluster visualization.

Passing the mouse over an event would reveal its title and offer the option to receive more information about the event or find similar events. Selecting the option to find similar events would issue an IN-SPIRE query which would in turn generate an interactive galaxy visualization or a text list of events related to the event under analysis.

To provide multiply ways of accessing information, all major sections have been linked together in an intuitive manner so that the user can seamlessly move from one section to the next, drilling down to find more information about a specific event, or broadening the search to find and visualize similar events. Each event description has links to the map, the schedule of that day, similar events, events of the same type, and the visual exploration.

As an application, *InfoStar* is rich in features, but it is designed so that the most important ones were most visible, and the flow from one section to another is seamless. Lacking a strong user profile, the goal was to interleave a robust set of more traditional text lists and new visual exploration functionality. This would provide uses the opportunity to self-select a preferred information access technique based on the situation and environment. Thanks to the use of web standards and careful design, the application worked on all devices tested and even exceeded expectations at times.

Lessons Learned and Future Work

In developing the SC2004 *InfoStar* application, we endeavored to generally characterize the needs of conference attendees, content owners and organizers to provide an easy-to-use and stream-lined interface for managing and searching conference information on mobile devices. The ensuing system proved to be highly successful as it provided a first effective step toward extending visual analytics beyond the desktop paradigm through integration with traditional methods for organizing and accessing conference information.

Developing an application based on the *InfoStar* platform concept for thousands of knowledgeable users provided us with great feedback. The application was used in a variety of environmental situations ranging from sitting in a quiet hotel room to walking the show floor with all the noise and crowds. Understanding how visual analytic capabilities such as the interactive galaxy display were used is providing us with insights that help us assess the utility and challenges deriving from the deployment of visual analytics on mobile platforms.

We are currently using the feedback received at SC2004 to improve the *InfoStar* conference application for SC2005. One such improvement consists in increasing the amount of information conference attendees have access to through mobile electronic media. *InfoStar* gave access to a wealth of technical information and information that could not be accessed any other way. However, user feedback determined that access to additional information, such as locations of venues not on the showroom floor, should be made available to attendees. Just like their paper counterparts, electronic information tools need to strive to provide information on everything a user might need.

Perhaps the most important lesson learned from the SC2004 *InfoStar* application is that the primary research effort for moving visual analytics into mobile environments is to overcome bottlenecks that stop electronic access to information from being the preferred choice of users. Providing unique information is obviously a good approach, but easy access to information could be even more critical. For example, offering an efficient interface is important, but if it takes the user longer to turn on their laptop or PDA than it does to pull out a printed book or ask a question of someone walking by then the interface may never be used.

Observations during SC2004 have driven us to focus more attention toward always-on devices such as cell phones or phone/PDAs. Always-on devices have limited screen real-estate and in some cases reduced interaction support (e.g. lack of pen-

based input) but have been designed for quick access. We are currently revising the *InfoStar* interface to take advantage of the immediacy of always-on devices while working around their limitations regarding display and interaction functionality.

Acknowledgements

References

Buchanan, G., Farrant, S., Jones, M., Thimbleby, Harold., Marsden, G., Pazzani, M., (2001) Improving Mobie Internet Usability. *10th International World Wide Web Conference*, May 2001, Hong Kong.

Campbell, C., and Tarasewich, P. Creating Notification (2004) Cues for Mobile Response Teams. *Proceedings of the 2004 DSI Annual Meeting*.

Ericsson, T., Chincholle, D., Goldstein, M., (2001) Both the device and the service influence WAP usability. In: Vanderdonckt, J., Blandford, A., Derycke, A. (Eds.), *Proceedings of IHM-HCI2001, Usability in Practice, vol. II.* Ce´padue`s-Editions, Toulouse, France, pp. 79–85.

Stoop, E. (2003) Mobility: Freedom of Body and Mind. In Emile Aarts & Stefano Marzano (eds.) *The New Everyday: Views on Ambient Intelligence*. 010 Publishers, Rotterdam, The Netherlands.

Hetzler, E., and Turner, A., (2004), Analysis Experiences Using Information Visualization. *IEEE Computer Graphics and Applications*, Sept/Oct 2004, pp 22-26.

Jones, M., Buchanan G., Thimbleby, H., Marsden, G., (2000) User Interfaces for Mobile Web Devices WWW9 Mobile Workshop Position Paper. 9th International World Wide Web Conference, May 2000, Amsterdam, The Netherlands.

Lee Y,E., and Benbasat, I., (2003) Interface Design for Mobile Commerce. *Communications of the ACM*, December 2003, pp 49-52.

Rex, B., Risch, J., Dowson, S., Moon, B., (1999), Multiple source information analysis, GIS and Starlight. *Proceedings of the 7th ACM international symposium on Advances in geographic information systems*, Kansas City, US., pp 158-159.

Risch, J.S., Rex, D.B., Dowson, S.T., Walters, T.B., May, R.A., and Moon, B.D. (1997). The STARLIGHT Information Visualization System. In *Proceedings of the 1997 IEEE International Information Visualization Conference (IV '97)*, August 27-29, 1997, London, England.

Tarasewich, P. (2003) Designing Mobile Commerce Applications. *Communications of the ACM*. Vol. 46/12, 57-60.