Interactive Visualization of Multivariate Statistical Data

Jim X. Chen, Daniel B. Carr, Harry Wechsler and Zhigeng Pan

Abstract—**This paper introduces web-based interactive Linked Micromap (LM) plots, a set of dynamic visualization methods that allows readers to interactively select variables and modify the different views to help reveal relationships among the study units. This methodology provided the foundation for web-based micromaps used by the National Cancer Institute (NCI). This illustrates the power of visualization to make statistical summaries involving health and risk factors for millions of people accessible to health planners than may have never had a statistics class. LM plots methodology is in use by the Department of Agriculture and readily extend to other application in other agencies in the United States. The interactive methods can be as useful in such extensions as they were for the National Cancer Institute.**

Index Terms— **Interactive visualization, Linked Micromap (LM) plots, statistical data display, web-based visualization.**

I. INTRODUCTION

Linked Micromap (LM) plots constitute a new template for the display of spatially indexed statistical summaries [1, 2]. This template has four key features: 1) displaying at least three parallel sequences of panel types that include micromaps, labels, and statistical summaries, 2) sorting of study units, 3) partitioning of the study units into perceptual grouping panels to focus attention on a few units at a time, and 4) positional linking of perceptual grouping panels across panel types and linking of study units across panel types typically using color and often using position.

Static LM plots have been used to visualize data sets varying size, complex, and domains [3, 4, 5]. The first effort toward a web application involved hundreds of hazardous air pollutants and estimates available for the US states, counties, and even census tracts [6, 7, 8]. This Environmental Protection Agency (EPA) research was stopped before releasing the web site to the public due to concerns about data quality and public reaction. The LM methodology introduced below is new except it utilized general map boundaries for the EPA funded research.

This paper introduces *web-based interactive* LM plots, a set of dynamic LM visualization methods that allows readers to interactively select variables and modify the different views to

help reveal relationships among the study units. This methodology provided the foundation for web-based micromaps used by the National Cancer Institute (NCI) [9, 10]. NCI uses the State Cancer Profiles web site to communicate with health planners across the nation. This illustrates the power of visualization to make statistical summaries involving health and risk factors for millions of people accessible to health planners than may have never had a statistics class. The statistical summaries in LM plots scale to large data sets and usability assessment confirms that they work for communication to health planners. LM plots methodology is in use by the Department of Agriculture and readily extend to other application in other agencies such as the Department of Homeland Security. The interactive methods can be as useful in such extensions as they were for the National Cancer Institute.

We use cancer statistics from the National Cancer Institute (NCI) as an application and implementation example to present the methods. The displays are of test data, not of official cancer statistics, but nonetheless provide an excellent test-bed for studying statistical visualization methodology. We have implemented a full-fledged set of LM plots for recent cancer statistical summaries of the United States at the state and county level. The test-bed list of selectable cancer sites is restricted to breast, colon, prostate, and lung, but readily extended. These web-based interactive LM plots have preserved all the key features of the LM plots originally published. While spatial resolution is lost in the transition from the printed page to a computer monitor, the interactive viewing options, allowed better visualization through drill-down views, multiple levels of detail, sorting, magnified micromaps, miniature overall statistical summary, confidence interval switching, and other interactive visualization methods. While the interactive methods are not new individually, their integration with LM plots provides a new approach to communicating spatially indexed statistical summaries over the Internet. Since the Internet is a widely accessible source of public information, the web based implementation of LM plots will make information available to more readers while the new interactivity can lead to more involvement and better understanding.

II. THE STRUCTURE AND FEATURES OF THE WEB-BASED INTERACTIVE LM PLOTS

Displaying LM plots on the Internet introduces a new and

effective way of visualizing various statistical summaries. Through only a web browser, a reader can easily access and view the statistical data in LM plots everywhere around the world. Effective web-based LM plots must include many interactions that require retrieving new data from the web server and displaying the retrieved data in different layouts to facilitate the readers. Java programming environment provides mechanisms that allow users to control the display content in a web browser interactively. In our implementations, we chose Java as the programming tool to achieve the web-based interactive LM plots of the national cancer statistical summaries.

In this section, we will introduce the structure and features of the web-based LM plots. A demo system with all the features is available on the website at http://cs.gmu.edu/~jchen/cancer/. It is much easier to understand the descriptions in this paper if the demo system is tried interactively along with the descriptions. This demo was developed using Internet Explorer and is best if viewed with that browser.

A.   Display panels and study units

The LM plots for displaying the cancer summary statistics have four parallel sequences of display panels (columns), which are US/State micromap, State/County name, and two cancer statistical summaries. Typically one summary concerns cancer mortality rates and the other concerns a cancer risk factor.

The study units (rows) are States or Counties that are interactively selectable depending on the current display options from a pull-down menu. For national cancer statistics, the study units are States. For state cancer statistics, the study units are Counties. The number of study units dynamically changes with current study units data. The number of the study units viewable in one frame at a time is about 41. When the actual number of study units is over this maximum, the whole display panel becomes scrollable. Fig. 1 and Fig. 2 are two snapshots showing the basic layouts of the cancer statistical LM plots of the United States and the state Kentucky, respectively.

In a row of a linked study unit, the geographic location, the leading dot before the name, and the statistics are all in the same color. On each statistics panel, the corresponding statistical value is shown as a dot with the confidence interval (CI) or bound displayed as line segments on both sides of the dot. The default scaling is set to include the lowest lower bound and largest upper bound of the CI's. However, when a cancer is rare and the population is small, the upper bound of the CI may be extremely large. The CI button is an icon composed of a point with line segments on each side and is located at the top of each statistical display panel. This button is used to switch between the default scaling that allows a complete CI view and a view that is scaled to the values of the maximum and minimum point estimates truncating the long CI's.

B.   Sorting and grouping the study units

Sorting the study units is an important feature of LM plots.

When sorting on a statistical variable, the displayed dot curve can show the relative relationships among the study units clearly. Here we allow interactive sorting on the study unit names and statistical variables panels to provide fast visualization from several different interchangeable perspectives. A sorting button with a triangle icon is located in front of the corresponding panel title. Click on a sorting button, causes the study units to be sorted according to the corresponding names or statistical variables. An up-triangle icon on the button represents ascending sorting, and a down-triangle icon shows a descending sorting.
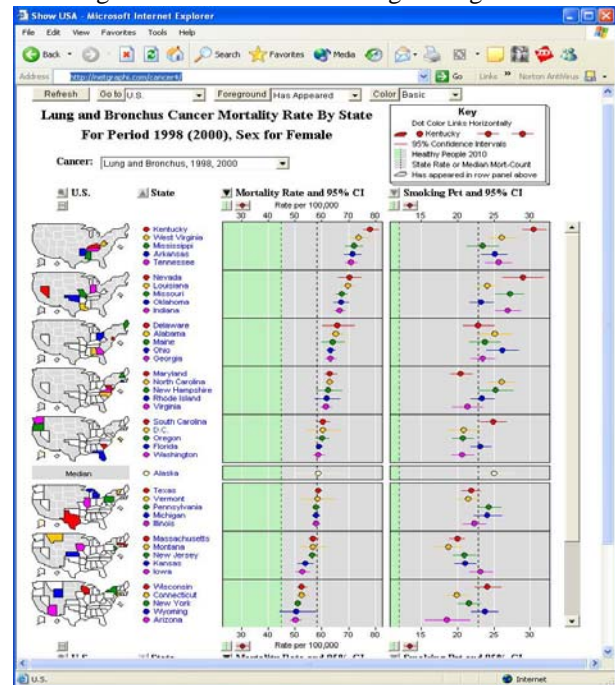

Fig. 1: A Snapshot of the LM plots of the U.S. Cancer Statistics See Color Plate 24.
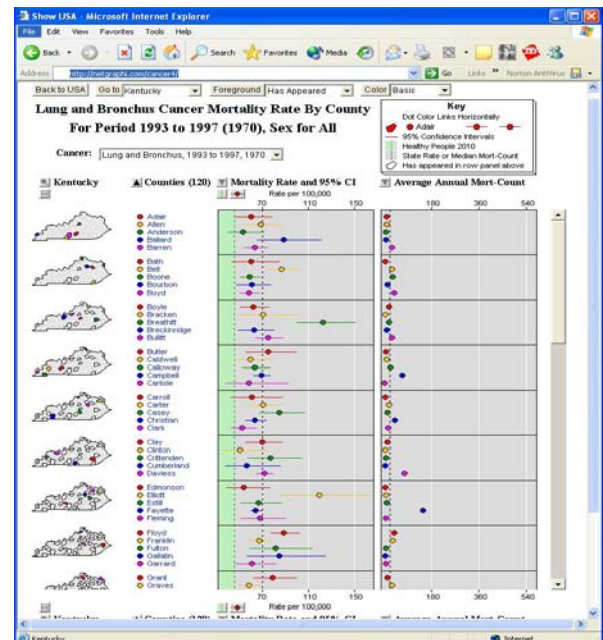

Fig. 2: A Snapshot of the LM plots of the Kentucky Statistics. See Color Plate 25.

For example, Fig. 1 shows that the study units (States) are sorted by one statistical summary (Cancer Mortality Rate) in the descending order. Fig. 2 shows that the study units (Counties) are sorted by the county names in the ascending order.

The study units are partitioned into a number of sub-panels vertically. This helps to focus attention on a few units at a time. We group every five consecutive study units into a sub-panel. In each group, a coloring scheme of five different colors is assigned to the five study units respectively. This can show the linked elements of one study unit within a group more clearly.

C.   Linking the related elements of a study unit

Linking all the elements of one study unit together helps to highlight a study unit. All the elements across the corresponding panels of the sequences are represented by the same color. Moreover, to highlight one study unit, the reader moves the mouse cursor onto any one of the related elements (micromap, unit name, or colored dots), and this causes all the linked elements in a study unit to blink. Meanwhile, the corresponding statistical summaries are displayed in text at the bottom of the browser window (in the browser's status line).

The coloring scheme for micromaps serves multiple objectives [4]. The basic objective is to facilitate rapid location of a particular study unit in a sub-panel. In our LM plots, the coloring scheme is repeatedly used for all the groups. For each group, a Sequential coloring scheme can help show the sequence of the study units in that group. In our implementation, a pull-down menu is provided for a reader to interactively select a different coloring scheme. The current design includes three coloring schemes: Spectral, Sequential, and Divergent. The default Spectral coloring scheme will work well for most users of the system. The Sequential coloring scheme is preferred by persons who are colorblind and for black and white printing. The Divergent color scheme will also work well for the colorblind user but is not suitable for black and white printing.

D.   Micromap and its magnification

Micromaps are the most active components in LM plots. Through the micromaps, a reader can find the locations of the study units intuitively. Furthermore, a reader can see the geographical distributions of the related study units in each group because the grouped study units are highlighted with the current color scheme in a corresponding micromap. When a reader moves the mouse onto a study unit in the micromap, the corresponding area will also blink with other linked elements.

In a micromap, except for the colored study units, all other study units are divided into two categories, which are called the background study units and the foreground study units. The background study units are displayed in a light gray color with a white outline. The foreground study units are displayed in white or light yellow according to the selected coloring scheme and outlined in black. The black outline overwrites a shared boundary that is white. A pull-down menu allows the user to control which study units belong to the foreground (or background). In the current design, the menu has the following

three choices:

1) Has appeared: Based on the sort order, the foreground study units in white or yellow are those that have appeared higher in the sort order and the background study units in gray are those that are lower in the sort order. The foreground coloring has two uses. First, it provides an index for the user searching the micromaps for a particular area of interest based on geography. If the area is in the foreground color, then it appears in a micromap group above in the sort order. If the area of interest is in the background color then it appears in a micromap group below. Second, the foreground color provides a cumulative geographic view of the ordered data that may show interesting geographic patterns such as a clustering of higher rates.

2) Will appear: This reverses what was seen in "Has appeared" so that the foreground study units in yellow or white are those that are lower in the sort order so have not appeared in previous micromaps. Similarly, the background study units in gray are those that are higher in the sort order so have appeared.

3) Above/below Median: For the micromaps above the median unit, foreground study units in white or yellow are those that are above the median unit, and the background study units in gray are those that are below the median unit. For the micromaps below the median unit, the meanings of the foreground and background study units are reversed.
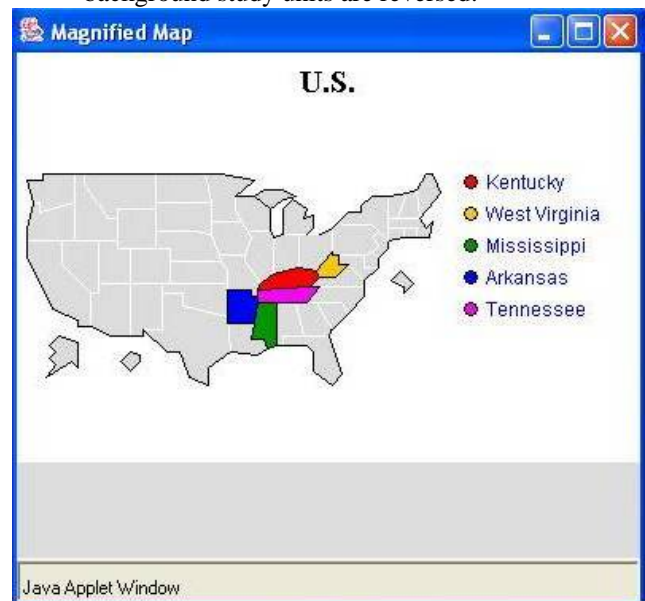

Fig. 3: The Magnified Micromap Window. See Color Plate 26.

Some states may include many counties (study units). For example, Texas has over 200 counties. Since each micromap occupies a small fixed area, the area of a study unit in a micromap may be too small for a reader to see its shape. Magnifying the micromap helps the reader to view all study

units clearly. We pop-up a new window and magnify the micromap in which the current highlighted study unit is located. Moving the mouse over a study unit on the magnified micromap causes the micromap in which the study unit is highlighted to appear in the magnified view. Fig. 3 shows a magnified U.S. micromap. Clicking the magnification button in front of the micromap panel title causes the pop-up magnified micromap window to appear as does right clicking either on a study unit name or on a dot in the corresponding statistical summary panels. Adjusting the size of the magnified micromap is handled by dragging the window to change the window's size.

E.   Automatic scrolling and navigation

Because the number of the study units in one LM plot may be larger than the maximum that the monitor screen can hold, some study units may not be displayed in the panel currently. Therefore, when a reader highlights a study unit from a micromap or from a magnified micromap and the study unit happens to be not in the display panel, the reader will not see the corresponding elements of the unit blinking. In our implementation, when this happens, the reader can click the right mouse button, and the display panel will automatically scroll to the exact position so that the corresponding study unit appears in the display panel. This operation can also help the reader rapidly locate on a study unit.

NCI usability studies indicated usability problems when both a windows scroll bar and a studies unit scroll bar appeared as in Fig. 1. We provided a solution to NCI that reduced scrollable area height context. Then reference information above and below scrollable area always remained visible and the windows scroll bar did not appear. This approach is now used in the NCI web site.

Drill-down is an operation that zooms in from a high-level LM plot to the corresponding low-level LM plot to view the detailed statistical data and the relationships among different hierarchical levels of detail. When using LM plots to visualize the national cancer statistical summaries, the U.S. LM plot provides a starting place for subsequent drill-down to any other state LM plot. In the U.S. LM plot, the state names and all the micromap regions are the active links to drill down to the corresponding state LM plots. In general, when moving the mouse cursor onto an element, if the mouse cursor becomes a hand-cursor, the element is then an active link. Clicking the left mouse button on that element drills down to the next LM plot level.

Drill-down is also available from a magnified micromap. When one area in the magnified micromap is blinking and the shape of the cursor changes to a hand with a pointing finger, clicking the left mouse button will also carry out the drill-down operation.

Navigation allows access to all the LM plots. Drill down is a special type of navigation. Our system also allows navigating from one LM plot directly to any other LM plot no matter what current LM plot level is. This is achieved through a pull-down navigation menu. From the menu, a reader can directly access the U.S. LM plot or any other state LM plot.

F.   Overall look and displaying different data sets

As we mentioned previously, when the number of the study units is bigger than the maximum that the display panel can hold, some study units will not be in the display panel. At this time, the reader cannot see the whole statistical curves formed by the dots shown in the statistical panels. However, the overall look of the whole pattern presents very useful information. We have added a pop-up window to display the pattern in a scaled down fashion. Fig. 4 shows such an overall look display of the miniature statistical summary. Clicking the button below the magnification button activates this display window.
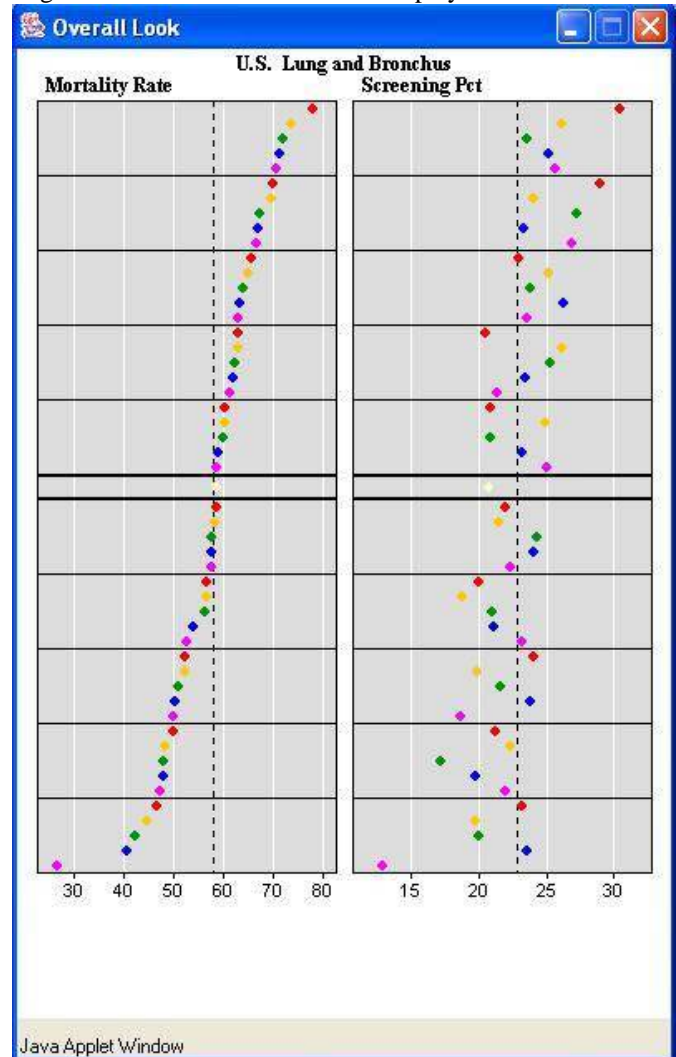

Fig. 4: Overall Look of the Statistical Panels. See Color Plate 27.

Interactively displaying different statistical data allows a reader to view and compare the relationships among different statistical results. In our system, we provide a pull-down menu for a reader to choose and view several different statistical cancer data sets. Once a reader selects a new cancer type, the data will be downloaded from the web-server and displayed in the display panel with a new calculated scale. In order to improve the display efficiency, the downloaded cancer data will be buffered in memory so that next time a reader selects

this cancer type again, the data can be directly fetched from the memory, instead of downloading from the web-server again.

III.   TECHNICAL PROBLEMS

In this section, we discuss our innovative solutions to address two technical problems and increase efficiency and utilization of the web-based interactive LM plots.

A.   Stateless programming environment

As we know, a web programming environment is stateless, which means that the states in the current web-page cannot be brought to the next web-page. For the web-based LM plots, each LM plot is actually a web page. When navigating among LM plots, a reader is actually accessing different web pages. As we mentioned above, web-based LM plots provide high interactivity. When interactively setting up some display features for a page, a reader generally hopes these setups can be brought to other pages by default. He doesn't want to set them again for each new page. Therefore, web-based LM plots must preserve the current setup states, just like a stand alone application that keeps all the states for the duration of the session. In other words, once a reader accesses an LM plot, all the states selected for that LM plot should be preserved to serve the whole session no matter where the reader navigates among different LM plots.

Generally, two methods can solve this problem. The first is to save the states in the request itself sent to the web-server. The second is to directly save the states at the web-server through some web programming tools like Microsoft ASP. Both of these methods require writing special code at the web-server side. This not only increases the programming complexity, but also affects the web accessing efficiency. Moreover, this may limit the web-server platform's independence from the web page contents. For example, Microsoft web-server supports ASP, but the web-based LM plots developed under this environment may not work for other web-servers.

Now, since we use Java as the programming language to develop the web-based LM plots, we can exploit Java's powerful programming capabilities to solve this problem in a new way. First, let's see why a web programming environment is stateless. The main reason is that the web-browser always uses http protocol to request html messages from the web-server. Unfortunately, http itself is a stateless protocol, and html is also a stateless language. A web-browser always gives up all the information in the previous web-page when it displays a requested message (html file) from the web-server. Under such a condition, when a web-page (actually the embedded code) is not allowed to write any states to the web-client's devices, the program code embedded in one web-page will not be able to directly pass their states to the code in another web-page if there is no special code that works on the web-server.

From the above discussion, we can see that the solution has to be in the program code embedded within one web-page instead of separate web-pages. If the program code embedded

in a web-page can directly request messages from the web-server and control the message display in the browser's window itself, the program states saved in the memory by the code will be available as long as the program session is alive. It is possible to keep the states from being flushed out of the web-browser because the returned messages are received by the program code instead of the web-browser.

When we use Java as the web programming tool, a Java applet can be designed to implement the above mechanism. Once a web-page is displayed by the web-browser, the embedded Java applet will take control to display the LM plots and to accept any user input actions located in the applet panel. If the applet detects that one action needs to retrieve new data from the web-server, it activates the communication code to request certain data from the web-server, and then displays the data in the applet panel. The result looks like a new "web-page" is displayed, but the fact is that it is just a programmed display refreshment. The newly displayed pages are not necessarily html pages. Instead, they can just be displays on the same Java applet panel that are painted by the same Java applet code according to the retrieved data. In application, this process is actually transparent to the web readers. So, a reader always feels that a new web-page is displayed when he clicks on an active link. (The reader may notice the discrepancy when he/she use "Back" or "Reload" in a web-browser that really invokes a different web-page.)

In our application, when a reader accesses an LM plot through a URL, the corresponding html page will be displayed by the web-browser. Meanwhile, the Java applet embedded in the html page will be activated. Then, the applet will download the default map data and cancer statistical data (generally is for the U.S.) from the web-server, and will finally display the corresponding LM plot on the applet panel, that looks like a web-page. This page actually is the starting point to view other LM plots. The next time, as long as the reader accesses other LM plots through the active links displayed in the current LM plot, the applet will repeat the above process to download new map data and cancer data from web-server and display a new LM plot again. Unless a new html page is requested by the web-browser, the above process will keep running (as an active session) and the applet will control all the operations to display different LM plots interactively according to the reader's input actions.

B.   Statistical data retrieval

In a statistical visualization application, efficient retrieval of the statistical data is very important. In a statistical visualization system, the statistical data may be saved in different formats, places or modes to expedite retrieval.

Firstly, the data format means the organization of the data flow received by the Java applet. The data format could be in binary, text, or some other format. The data format is generally not a problem for Java code. Actually, a Java code can handle any data format as long as it has certain rule to be followed.

Secondly, the statistical data can be saved anywhere. It may be on the same system of the web-server, or on a different

remote system. No matter where the system is located, as long as the system is connected to the Internet, a Java code can always retrieve it through the TCP/IP protocol.

Finally, the statistical data saving mode needs to be addressed carefully. Currently, there are two general data saving modes: file or database. If the data is saved in the file mode, as long as the read access right is open, a Java code can directly read the data file through the Internet. However, if the data is saved in a database system, the situation becomes complex. Generally speaking, through the JDBC (Java Database Connectivity) technology, a Java code can retrieve the data saved in a database. However, for security reasons, a safety database system is usually not open to the public. Therefore, the Java applet embedded in a web-page is generally not allowed to access a database system through JDBC directly. We believe the way to access the data is through the web-server, to which the database system may release the access right. Under such a situation, if a Java applet wants to retrieve the data from the database, it has to get the aid from the web-server. This means that a data retrieval code with the JDBC API has to be written on the web-server side. In this way, when a Java applet needs to retrieve a data from a database, it first tells the data retrieval code on the web-server what data is needed; then, the web-server code connects to the database system, retrieves the data from the database through the JDBC API, and sends the retrieved data back to the applet code at the web-client side; finally, the applet formats the received data and displays the plots on the applet panel. To provide a high efficiency, the data retrieval code on the web-server can be coded in Java Servlets.

In our application, the statistical data is saved in the file mode and on the web-server. We don't need to write any code for the web-server. Instead, the Java applet directly retrieves the cancer data from the files.

IV. CONCLUSION

LM plots are being increasingly used. An increase number of tools are available for the production of static displays. This includes an NCI Java application develop to help screen cancer registry data for data quality. The NCI Java applet source code is available to other federal agencies and has been available to individual requesting it. Students at George Mason University modified the NCI data base query portion of the NCI applet to access EPA's toxic release inventory over the web and produce corresponding county level LM plots. Research at Utah States modified the NCI applet for showing West Nile Virus patterns [11]. New applications continue to emerge in France, Japan, and other countries [12, 13, 14].

We have described the design and implementation of a set of web-based interactive LM plots, a statistical data visualization system that integrates geographical data manipulation, visualization, interactive statistical graphics, and web-based Java technologies. The system is effective in presenting the complex and large-volume sample data on the national cancer statistics. With some modifications, the web-based interactive

LM plots can be easily applied to visualize many other spatially indexed statistical datasets over the Internet. We believe our system is unique in its design and integration with the creative methods mentioned in the paper. Creative research will find a host of application and likely some useful extensions.

REFERENCES

[1] D. B. Carr and S. M. Pierson. Emphasizing Statistical Summaries and Showing Spatial Context with Micromaps, *Statistical Computing & Statistical Graphics Newsletter*, vol. 7, no. 3, pp. 16-23, 1996,
[2] D.B. Carr, A.R. Olsen and et al. Linked Micromap Plots: Named and Described, *Statistical Computing & Statistical Graphics Newsletter*, vol. 9, no. 1, pp. 24-31, Summer 98.
[3] D.B. Carr, A.R. Olsen, S.M. Pierson and J.P. Courbois. Using Linked Micromap Plots to Characterize Omernik Ecoregions, *Data Mining and Knowledge Discovery*, vol. 4, pp. 43-67, 2000.
[4] D. B. Carr. Designing Linked Micromap Plots for States with Many Counties, *Statistics in Medicine*, vol. 20, pp. 1331-1339, 2001.
[5] D. B. Carr, J. F. Wallin and D. A. Carr. Two New Templates for Epidemiology Applications: Linked Micromap Plots and Conditioned Choropleth Maps, *Statistics in Medicine*, vol. 19, pp. 2521-2538, 2000.
[6] J. Symanzik, D. Wong, J. Wang, D. B. Carr, T. J. Woodruff and D.A. Axelrad. Web–based Access and Visualization of Hazardous Air Pollutants, *Geographic Information Systems in Public Health: Proceedings of the Third National Conference of Agency for Toxic Substances and Disease Registry,* August, 1998.
[7] J. Symanzik, D .A. Axelrad, D. B. Carr, J. Wang, D. Wong and T. J. Woodruff. HAPs, Micromaps and GPL — Visualization of Geographically Referenced Statistical Summaries on the World Wide Web, *Proceedings of ACSM–WFPS–PLSO–LSAW Conference (CD), American Congress on Surveying and Mapping*, 1999.
[8] J. Symanzik, D. B. Carr, D. A. Axelrad, J. Wang, D. Wong and T.J. Woodruff. Interactive Tables and Maps — A Glance at EPA's Cumulative Exposure Project Web Page, *Proceedings of the Section on Statistical Graphics of American Statistical Association*, pp. 94–99, 1999.
[9] D. B. Carr, J. Chen, S. Bell, L. Pickle and Y. Zhang. Interactive Linked Micromaps and Dynamically Conditioned Choropleth Maps, *Proceedings of the Second National Conference on Digital Government Research*, Digital Government Research Center, pp. 61-67, 2002.
[10] D. B. Carr, S. Bell, L. Pickle, Y. Zhang and Y. Li. The State Cancer Profiles Web Site and Extensions of Linked Micromap Plots and CCmaps, *Proceedings of the Third National Conference on Digital Government Research*, Digital Government Research Center, pp. 269-273, 2003.
[11] J. Symanzik, S. Gebreab, R. Gillies and J. Wilson. Visualizing the Spread of West Nile Virus, *Proceedings of American Statistical Association (CD)*, 2003.
[12] D. B. Carr. Some Recent Graphics Templates and Software for Showing Statistical Summaries, Eds. C. R. Rao, E. Wegman, and J. Solka, H*andbook of Statistics,* vol. 24, Data Mining and Visualization, North Holland, New York, pp. 415-436, 2005.
[13] S. Lee, T. Chen, J. Kim, G. J. Kim, S. Han and Z. G. Pan. Using Virtual Reality for Evaluating Affective Properties of Product Design, *IEEE VR 2004 Conference*, 2004.
[14] J. F. Lu, Z. G. Pan, etc. Virtual Learning Environment for Medical Education Based on VRML and VTK, *Computers &Graphics*, vol. 29, no. 2, pp. 231- 238, 2005.
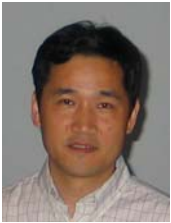
Jim X. Chen was born in Chengdu, China on August 6, 1962. He received his B.S. and M.S. in computer science from Southwest Jiaotong University, Chengdu, China in 1983 and 1986, respectively. He received his Ph.D. in computer science from the University of Central Florida in 1995.

He is an Associate Professor in the Department of Computer Science at George Mason University (GMU), the Director of the Graphics Lab at GMU, Associate Editor-in-Chief (EIC) of AIP/IEEE Computing in Science & Engineering (CiSE), and Associate EIC of International Journal of Virtual Reality (IJVR). He served as General Co-Chair of IEEE VR2006 and Guest Editor for IEEE Computational Science & Engineering, CiSE, and PRESENCE. He authored two books and two book chapters, edited two conference proceedings, published over 70 research papers, and acquired three patents. His research interests include computer graphics, virtual reality, visualization, networking, and simulation.

Dr. Chen is a member of IEEE. He received honorary professorships at Beijing Jiaotong University, Fudan University, Hoseo University, Sichuan Normal University, Southwest Jiaotong University, and Xihua University. Contact him at jchen@cs.gmu.edu.

Daniel B. Carr earned a B.A. in mathematics and psychology from Whitman College in 1968, an M.Ed. in counseling from Idaho State University in 1972, an M.S. in statistics from Oregon State University in 1972, and a Ph.D. in statistics from the University of Wisconsin, Madison, in 1976.

He is Professor of applied and engineering statistics in the Department of Applied and Engineering Statistics at George Mason University. Prior to joining George Mason University he worked as a senior research scientist and technical working group leader of the Exploratory Data Analysis Group in the Computational Science Department of Battelle Pacific Northwest Laboratories. He has written over 65 papers with many developing graphical and computational methods for exploratory analysis of large data sets. His research addresses many of the challenges in statistical graphics design, including representation of large data sets, visual data mining, and web communication.

Dr. Carr is a Fellow of the American Statistical Association and a columnist for the joint newsletter of the Statistical Computing and Statistical Graphics sections. Contact him at dcarr@gmu.edu.

Harry Wechsler received the Ph.D. in computer science from the University of California, Irvine, in 1975.

He is a Professor of computer science and Director for the Center of Distributed and Intelligent Computation at George Mason University (GMU). He has published more than 250 scientific papers, serves on the editorial board of major scientific publications, and is the author of Computational Vision (Academic Press, 1990). As a leading researcher in face recognition, he organized and directed the NATO Advanced Study Institute (ASI) on "Face Recognition: From Theory to Applications" (Stirling, UK, 1997), whose seminal proceedings were published by Springer (1998). His book on Reliable Face Recognition Methods, which breaks new ground in applied modern pattern recognition and biometrics, will be published by Springer in the fall of 2006. His research in the field of intelligent systems focuses on computational vision, image and signal processing, data mining, machine learning and pattern recognition, with applications for ATR, biometrics/face recognition, augmented cognition and intelligent HCI, change detection and link analysis, performance evaluation, and video processing and surveillance.

Dr. Wechsler was elected an IEEE Fellow in 1992 for "contributions to spatial/spectral image representations and neural networks and their theoretical integration and application to human and machine perception;" and an IAPR (International Association of Pattern Recognition) Fellow in 1998. Contact him at wechsler@cs.gmu.edu.

Zhigeng Pan received his BS and MS from the Computer Science Department at Nanjing University respectively in 1987 and 1990, respectively, and his PhD from Zhejiang University in 1993.

Since 1993, he has been working at the State Key Lab of CAD&CG on a number of academic and industrial projects related to distributed graphics, virtual reality, and multimedia. He has published more than 120 papers in international journals, national journals, and international conferences. He is the author or co-author of four books related to computer graphics. He acted as one of the Guest Editors of a Special Issue of Computers & Graphics ("Computer Graphics in China"). Currently, he is the Editor-in-Chief of The International Journal of Virtual Reality. He is on the editorial board of International Journal of Image and Graphics, International Journal of CAD/CAM, Journal of Image and Graphics, Journal of CAD/CG, et al. He is the Chairman of the Virtual Reality Committee, China Society of Image and Graphics; the Depute Director of the Machine Perception and Virtual Reality Committee, China Association of Artificial Intelligence; the Steering Committee member of VRCAI series conference. He has been the Committee Co-Chair of VRAI'2002, VRAI'2003, ICIG'2002, Eurographics/ACM SIGGRAPH Workshop on Computer Graphics Education'2004, and ICIG'2004. He is the Program Co-Chair of EGMM'2004 (Eurographics workshop and ICIG'2005 on Multimedia), Edutainment'2005 and Edutainment'2006. He also served as Conference Co-Chair of ICAT'2006.

Dr. Pan is a member of SIGGRAPH, Asia reprehensive of SIGGRAPH Education Committee，Eurographics, and IEEE. He is also a senior member of the China Image and Graphics Association. He is on the director board of the International Society of VSMM (Virtual System and Multimedia), a member of IFIP Technical Committee on Entertainment Computing (acting as representative from China).