TECHNICAL RESEARCH REPORT

Dynamic Query Chloropleth Maps for Information Seeking and Decision Making (2003)

by Kent L. Norman, Haixia Zhao, Ben Shneiderman, EvanGolub

TR 2005-43

Dynamic Query Chloropleth Maps for Information Seeking and Decision Making

Kent L. Norman

Department of Psychology
University of Maryland
College Park, MD 20742
kent_norman@lap.umd.edu

Haixia Zhao, Ben Shneiderman, and EvanGolub

Human-Computer Interaction Laboratory
University of Maryland
College Park, MD 20742
haixia@cs.umd.edu, ben@cs.umd.edu,
egolub@cs.umd.edu

Abstract

Information retrieval and visualization can be combined in dynamic query systems that allow users unparalleled access to information for decision making. In this paper, we report on the development and evaluation of a dynamic query system (YMap) that displays information on a chloropleth map using double thumb sliders to select ranges of query variables. The YMap prototype is a Java-Applet that supports panning and zooming. Several usability studies were conducted on early prototypes that resulted in the current version. Applications of YMap for decision making tasks are discussed.

1 Introduction

New methods of information visualization open up rich databases to information analysts by taking advantage of the power of human visual processing in combination with complex screen displays (Card, MacKinlay, & Shneiderman, 1999). Such visualizations are particularly important in mapping geographic data (MacEachren, Hardisty, & Gahegan, 2001). Map displays can combine a number of facets of nominal and quantitative data to show vast amounts of data in a way that illuminate rather than obscure important points and trends. For example, choropleth maps combine geographic layout with quantitative information by colourizing areas using a colour gradient. Choropleth maps allow analysts to quickly scan the area and to locate highs, lows, and trends (Andrienko & Andrienko, 1999).

Similarly, new methods of dynamic database query facilitate information retrieval by creating a close coupling between the direct manipulation of query variables and the presentation of the retrieved set (Ahlberg & Shneiderman, 1993). Dynamic query applications allow the users to see the immediate result of changes in the query variables.

This paper presents a new prototype of a dynamic map query system called "YMap" based on the work of Dang, North, and Shneiderman (2001).

2 YMap Prototype

Figure 1 shows a screenshot of YMap applet. It contains a choropleth map (centre/left), a set of dynamic query double-box sliders (right/middle), a detail data panel (bottom/left), a drop-down-list-box of attributes to choose from for shading the map (above the map, with continuous shading legend), an overview map (right/bottom), and a toolbox for panning, zooming and toggling between a state view and a county view (top). The prototype follows the general interface design methodology for exploring large data sets proposed by Shneiderman (1998): overview of the data distribution via visualization (shaded choropleth map and histogram bars on the sliders), dynamic query (via double-box sliders) to filter out unwanted entries and narrow down the result, and details-on-demand to further examine individual entries of interests (mouse over a region to view the data associated with that region, select regions to compare their data in the detail information panel, and save selected data to an external file for further analysis). The screenshot shows a county view of continental United States shaded by the year 1992 furniture and home furnishing store sale volume. Counties with low number of new private housing units or with median population age out of range [26, 42] are filtered out (in grey, the same colour used for out-of-range parts of the sliders.)

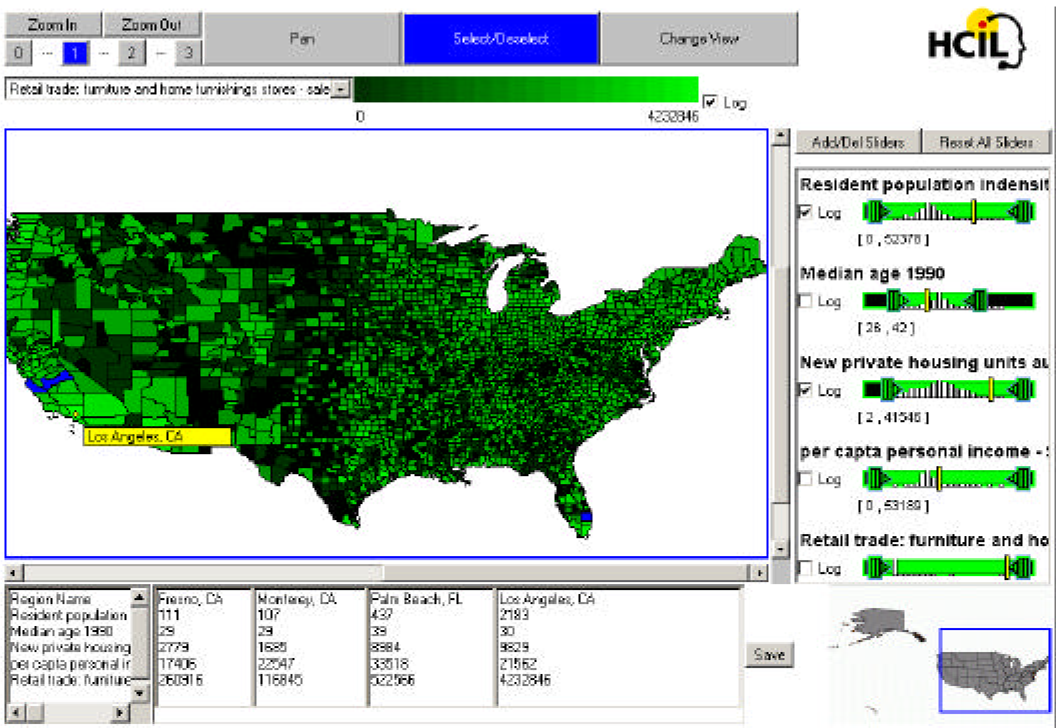


Figure 1: A screen shot of our YMap prototype

While a lot of additional features can be added to our prototype, such as scatter plots and linked brushing with the map, here we focus on three key issues in our system design:

2.1 Striving for universal usability

Our system strives to work towards universal usability (Hochheiser & Shneiderman 2001). Specifically, we aim to support occasional users with modem speed connections and low-end machines, as well as users with highspeed connections and fast machines. While long initial download time is a typical problem with most interactive Web map solutions that deliver vector-format map data to the browser for rendering, YMap employs an image-based technique that encodes geographic knowledge into raster images called base maps (Zhao & Shneiderman 2002).

This enables sub-second, fine-grained interface controls for dynamic query, dynamic classification, and geographic object data identification, with short initial system response time (to reduce the frustration of users with slow connections), near linear interaction performance scalability to the number of geographic objects (which accommodates users with low-end machines), and the potential of reducing the overall data transfer over the network.

2.2 Zooming/panning structure

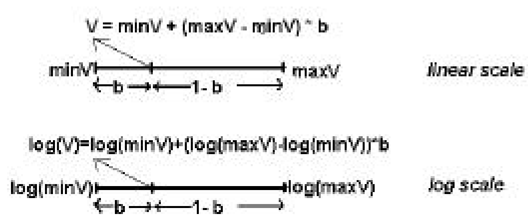
With the image-based technique, zooming/panning were designed to ensure usability as well as quick the system response time. We chose to pre-generate base maps and define the number of zooming levels instead of having the server generate the base maps on-the-fly and allowing free-scale zooming. Several strategies for achieving fast zooming/panning are discussed in by Zhao & Shneiderman (2002), such as breaking a large base map into several smaller ones, two-staged zooming response using aggregated/interpolated base maps, and pre-fetching.

2.3 Slider Structure

2.3.1 Data distribution and the histogram

While it is a good idea to present each geographic object as a dot along the slider (Andrienko & Andrienko, 1999), it is often hard to judge changes when a section of the slider is overcrowded (e.g., 20 dots and 200 dots of size 1mm^2 /ea along a 2mm slider section shows no difference to users' visual perceptions). YMap sliders use histograms (see Figure 2). The height of a histogram bar is proportional to the number of geographic objects with the attribute value falling in its range. When users mouse-over a bar, a pop-up appears showing the value range of the bar and its frequency (e.g., in Figure 2a, the highest bar shows that 8 states have resident population density within the range [62.76, 99.41] 1/SQM). The histogram serves several purposes: (a) it gives an overview of the statistical data distribution; (b) it gives users hints of how the map will be filtered when the slider is adjusted; and (c) it allow bi-directional linked identification between the map and the slider.

(a) Log scale. (b) Linear scale.



(c) Equations

Figure 2: Slider scales

When users mouse-over a region on the map, the value of that region is marked as a yellow bar on the slider (see Figure 3a). On the other hand, when users mouse-over a histogram bar, all the regions with the attribute values falling into that bar can be highlighted on the map (see Figure 3b) (to be implemented). In our prototype, users can move the two slider boxes to define a query width, then click a point between the two boxes, hold and move the fixed-width query range to view the change on the map (see Figure 3c).

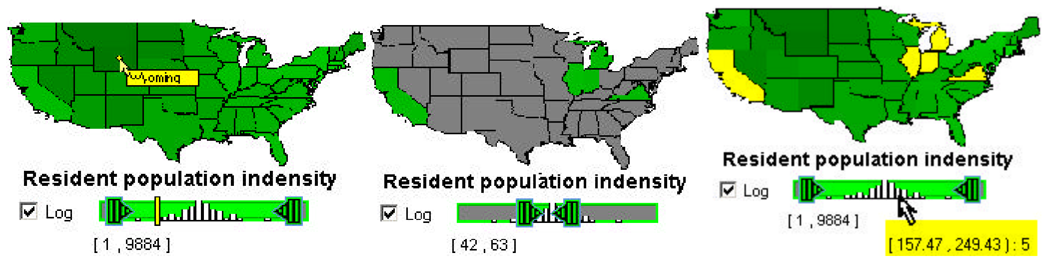


Figure 3: Bi-directional linked identification

2.3.2 Slider scales

Two different slider scales are supported, linear and logarithmic. In linear scale, the attribute value represented by a slider position is linearly proportional to its distance to the ends of the slider. In logarithmic scale, the values are logarithmized (see Figure 2c). The purpose of providing different scale options is to enable users to spread geographic objects along the slider, so filtering is more gradual when the slider is adjusted. As an illustration, in Figure 2b with a linear scale, 46 out of 51 states have their population density values falling into the first histogram bar. A slight movement of the slider box across the bar causes most of the states to be filtered out, hiding their difference from the exploration. When switched to logarithmic scale as in Figure 2a, states are more evenly distributed along the slider. The logarithmic scale used in our prototype helps to spread out overcrowded lower-end data. Other scale alternatives could be added, such as a percentile scale that would evenly distribute the geographic objects along the slider.

3 Usability

Several studies were conducted to evaluate the usability of earlier prototypes of the dynamic map query system. One of these is reported by Norman (2002). In this study, 43 undergraduate students from a psychology class were tested individually. First, they were instructed on how to set sliders to find groups of states on the U.S. map. Then they were given 18 problems using single sliders and 12 problems using from two to six sliders in combination. Finally, they were given a follow-up questionnaire on ease of use. Overall, subjects had little problem adjusting the single sliders and even complex combinations of the sliders to find sets of states whose values were within specified ranges. However, in addition to minor layout changes, these studies indicated the need to coordinate the colours of the sections of the sliders with the map areas that were selected versus not selected so as not to confuse which was which. Overall, the usability of YMap was rated fairly highly.

In addition, Norman (2002) found that with dynamic query maps, users quickly became aware of attributes that varied in geographic directions (e.g., north-south, east-west, coastal-central, etc.).

They also were able to detect when sliders were correlated with other sliders. Thus, dynamic query maps have the added benefit that users become more aware of the underlying structure of the database in the process of retrieval and exploration.

4 Applications

Dynamic choropleth query maps such as YMap have many applications in complex information analysis and decision making tasks. In addition to finding specific geographic regions that match a query and retrieving details, they are particularly usual in terms of sensing changes, gradients, and relationships among variables. They are particularly useful for making decisions among large sets of geographic alternatives. Decision with many alternatives generally involves a two-stage process: rapid elimination of unacceptable alternatives followed by a detailed comparison of remaining possibilities. YMap allows the user to home in on a select set of regions by adjusting sliders. For example, if one wanted to find the ideal region for a new manufacturing facility, the decision maker would eliminate many locations based on building cost or proximity to the market, select in locations with large labour forces, and select out areas with high costs of energy, etc. Thus, the first stage of the decision process involves a number of tradeoffs between the settings of the alternative sliders, adjusting one up and another down to reduce the set of alternatives to manageable, best set. The second stage involves a detailed comparison of the specific attributes of the alternatives. YMap provides the dynamic visualization interface to facilitate both stages.

References

Ahlberg, C., & Shneiderman, B. (1993). Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In R. M. Baeker, G. Grudin, W. A. S. Buxton, & S. Greenberg (Eds). *Readings in human-computer interaction: Toward the year 2000* (pp. 450-456). San Francisco: Morgan Kaufmann.

Andrienko, G. L. & Andrienko, N. V., (1999). Interactive maps for visual data exploration, *International Journal of Geographical Information Science*, 13 (4), 355-374.

Card, S. K., MacKinlay, J. D., & Shneiderman, B. (Eds.) (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufman.

Dang, G., North C. & Shneiderman B., (2001). Dynamic queries and brushing on choropleth maps, *Proc. International Conference on Information Visualization 2001*, IEEE Press, 757-764.

Hochheiser, H. & Shneiderman, B. (2001). Universal usability statements: Marking the trail for all users. *ACM Interactions*, 8 (2), 16-18.

MacEachren,A. M., Hardisty, F. & Gahegan, M. (2001, May). Supporting visual integration and analysis of geospatially-referenced data through web-deployable, cross-platform tools. Paper presented at the National Conference for Digital Government Research, Los Angeles, CA.

Norman, K. L. (2002). *Dynamic Query Maps: A Study of Performance and Usability*. (Report No. LAP-TR-2002-03). College Park, MD: Laboratory for Automation Psychology, University of Maryland.

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective human-computer interaction.* Reading, Mass: Addison-Wesley.

Zhao, H. & Shneiderman, B. (2002). *Image-based Highly Interactive Web Mapping for Geo-referenced Data Publishing*. (Report No. HCIL-TR-2002-26, CS-TR-4431, UMIACS-TR-2003-2). College Park, MD: Human-Computer Interaction Laboratory, University of Maryland.