

VizBlog: A Visualization Tool for Blog Discovery

Candida Tauro, Sameer Ahuja, Manuel A. Pérez-Quñones, Andrea Kavanaugh, Philip Isenhour
Center for Human Computer Interaction
Department of Computer Science, Virginia Tech
Blacksburg VA, USA 24061-0106
{ctauro, sameer.ahuja, mperezqu, kavan, isenhour}@vt.edu

ABSTRACT

Web logs (or blogs) have become a means for citizens to share opinions and deliberate on local issues. However, the large number of blogs makes finding and exploring content of interest relatively difficult. We have developed a tool that displays citizen-to citizen discussions in blogs and reveals similarities across blog entries. In this demo, we will show our tool in use with a digest of regional Southwest Virginia blogs. We will demo how users can navigate and explore online discussions by manipulating the graph, filtering content, and clicking on different parts of the graph.

• **INTRODUCTION**

Online deliberation is a term associated with an emerging body of practice and research dedicated to fostering purposeful discourse over the Internet. Weblogs can be used for online deliberation but are quite different from the traditional centralized discussion forums in terms of knowledge exchange patterns, the participants they attract, and the decentralized structure they employ. For example, the majority of bloggers are not political activists, but they do tend to be relatively well informed on a variety of topics and issues. In addition, the decentralized nature of blogs makes them easy to use because users can offer informal observations on various issues without the constraints of rules and formality associated with online forums. Instead of trying to find a centralized site where conversation is directed, bloggers set up their own sites (blogs), and just start writing about various topics (typically, “my life and experiences” although political opinions, observations and information are scattered throughout many of the “my life” kinds of blogs).

Ironically, these very characteristics that make blogs attractive to a broad and diverse set of voices, are also fundamental to the problems associated with using blogs for online deliberation. The overwhelming number of blogs full of a wide array of topics creates a discovery problem for users. To address this problem we present a tool to help find and participate in citizen-to-citizen discussion that takes place in blogs.

• **VISUALIZING BLOGS WITH VIZBLOG**

Our visualization tool, VizBlog (Figure 1), scans a predetermined set of blogs and creates a network visualization of citizen discussions. Through association and content analysis, blog entries are represented by nodes and linked to each other to form clusters of related local content. VizBlog has been developed using the Prefuse visualization toolkit¹. The visualization is made up of nodes and links. Figure 1 shows the main window for VizBlog. The right panel provides features for navigating, filtering and searching the visualization. These features are discussed below.

¹ <http://prefuse.org/>

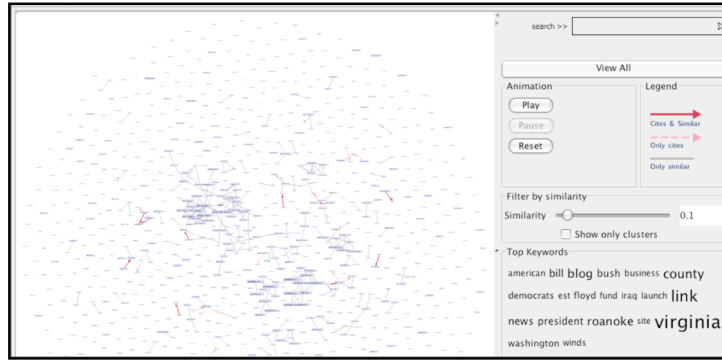


Figure 1. VizBlog’s User Interface

Each node in the visualization represents an individual blog entry, and is labeled with the title of the blog entry. The color of the nodes changes dynamically based on whether the node is selected, or is the neighbor of a selected node or if it is a search result. The size of a node is in direct proportion to the number of neighbors, i.e., the number of links it has. This form of coding stems from the observation that the blog entries that are central to a ‘conversation topic’ receive the most citations, and most of the peripheral blog entries in the conversation are similar to these central entries.

The links between the nodes represent relationships. Two nodes can be linked if there is a citation from one to another, or if they are similar in content as determined by the software, or both. Graphically, the three types of links are rendered differently in VizBlog to provide extra information to the user about the relationships between nodes.

The default view is an overview that provides a ‘starting point’ to the users so they can quickly recognize the major features of the visual space, and start exploration. A one month data set for a regional blog aggregator contains about 1000 blog entries. With such high number of nodes, cluttering and occlusion become significant problems at the overview levels. VizBlog addresses this problem by providing several ways of filtering and reducing clutter. Semantic zooming reduces the visual representation of nodes at overview levels by decreasing the length of the node title, further reducing clutter and occlusion (Figure 2). The tool also provides for zooming into areas of interest to the user. A user can right click-drag on any point on the display to zoom in or out with the display centered at that point. In addition, users can zoom into a particular area of interest by pressing the middle mouse button (or scroll-wheel) and drawing a rectangle over the area. Panning is supported via left-click drag on an empty area.

Figure 2: Semantic zooming – varying title length as the zoom value increases from views A to C.

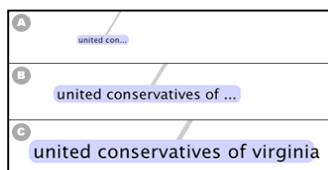


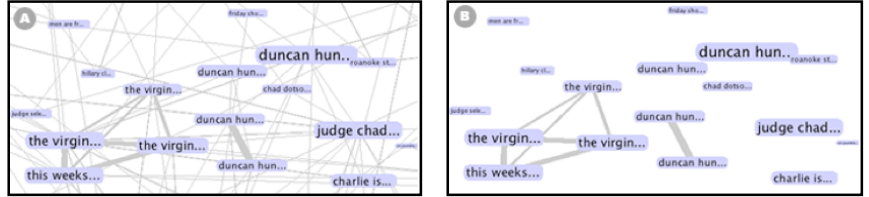
Figure 3: Top keywords cloud



The tool also displays the top keywords extracted by the preprocessor as a cloud in the right side panel. This cloud provides the viewer a depiction of the most used keywords in the data set (Figure 3). VizBlog also provides several features meant to aid filtering of data items in the visualization. For example, a similarity slider (shown in the middle of the right panel of Figure 1) lets the users filter edges based on their strength. This strength is determined by the value of the similarity coefficient for the two nodes connected by the edge. By moving a slider, the user can modify the cut-off value. This has an effect of

changing the density of the clusters. An example of this change is shown in Figure 4 below. Since we are interested in deliberation in the wild with an emphasis on citizen-to-citizen deliberation, blogs that are not connected to other entries are considered not relevant. Thus, the tool provides an option to filter off nodes that are not connected to any other node in the current view.

Figure 4: Using similarity slider to filter-off weak links. (A) shows all the links and (B) shows some links filtered out



The two filtering options significantly reduce the number of visible visual items, and increase the visibility of clusters. Figure 5 shows an example of how the user can quickly reduce the number of visual items to view only the most similar entries.

Figure 5: Identifying clusters is significantly easier after hiding unconnected nodes (B).

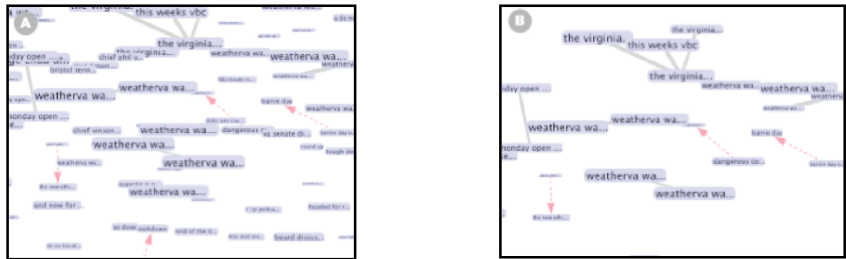


Figure 6: Mouse-over of a node reveals its title and highlights its neighbors

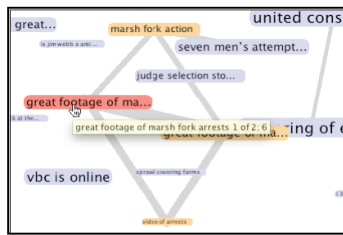
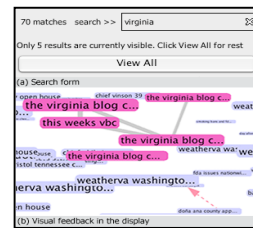


Figure 7: VizBlog search showing the results of searching for “Virginia”



Eventually, users would want to explore further details of particular entries. In VizBlog there are three ways to obtain more details as needed. First, right clicking on a node opens the particular blog entry in a browser window. Second, mouse over a node displays more information about the node (see Figure 6). Finally, searching highlights nodes that match the search term (see Figure 7).

We are grateful to the National Science Foundation Digital Government Program (IIS-0429274) for supporting the research described in this paper. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.