Pixel-Oriented Treemap for Multiple Displays

Haeyong Chung, Yong Ju Cho, Jessica Self, and Chris North

Virginia Tech

Abstract

We have developed a Pixel-oriented Treemap visualization intended for use on multiple displays with collaborating users. It visualizes the health and status of about a million devices with a Treemap layout. In this paper we describe how we found useful pieces of the VAST 2012 Challenge MC1dataset and discuss how users interacted with this visualization during the analysis.

Keywords: Large display, multiple displays, pixel-oriented visualization, treemap, physical navigation

Index Terms: H.5.2 [Information Systems]: Information Interfaces and Presentation—User Interfaces

1 Introduction

The data set for the VAST Challenge 2012 Mini Challenge 1 (MC1) requires a large scale situation awareness analysis to understand a large data set containing the network health and activity status of approximately one million online devices for three days. One of the tables in the dataset contains 158,530,955 records and the devices can be ordered by hierarchically based on business units and facilities. So the main visualization challenge was to support very large quantities of hierarchical information. To analyze such big data, we developed a pixel-oriented visualization with a Treemap layout [1], which is designed specifically for use on large, high-resolution displays, and the visualizations also allow the employment of personal devices including laptops, desktops, and even tablets as well as a large display to analyze data collaboratively.

2 The Tool Overview

The system consists of two main views: the Overview and Detail views. The Overview visualization on the large screen display allows multiple users to explore high level patterns through an aggregated view, while the Detail view enables users to examine the detailed information of each device through a coordinated pixel map on individual computers (Figure 1). The combination of the individual and shared visualization spaces allows for the distribution of analysis tasks among analysts thereby promoting improved efficiency in analyses of a large data set.

The Overview. The Overview visualization can represent the overall health and status of the entire BankWorld, such as policy status, activity flag, and number of connections within each region of BankWorld. As shown in Figure 2, its basic visual representation is based on Treemap in which each big rectangle represents a region (i.e. business units), and the online devices with unique IP addresses are represented by small squares. Each level of the policy status (Figure 2a) and activity flag (Figure 2b) are represented in 5 different colors (Figure 2e) and the number of connections are represented by Heatmap containing 100 levels (Figure 2c,f). The color black represents machines that are currently offline. This visualization is space-filling to maximize the use of every available pixel on the large display to visualize the entire health and status of devices at a specific time, so we could easily scale to a million devices fully utilizing the large display's high-resolution. For interaction methods, this visualization is designed to support physical navigation [2] but users can also change and select a Treemap for a specific date, time and data type using the mouse.

{chungh, ycho76, jazeitz, north}@vt.edu

Figure 1. A multiple display environment for Pixel-Oriented Treemap

The Detail View. The Detail View runs on the small displays and it enables users to zoom and examine one part of a region on the Overview Treemap in more detail (Figure 3). Users can explore specific areas of concern by breaking the big Treemap down into manageable chunks by selecting a date, time and region on the tooltip (Figure 3a). In this view, users can see detailed domain-specific information such as machine class, function, unit, facility, etc (Figure 3b). Users can hover their mouse over each pixel (colored squares, Figure 3c) in order to quickly browse the detailed information.

To facilitate a better insight into the changes and anomalies over time, we created an animated video with all snapshots of our Treemap visualization for three days. The video juxtaposes three different data including policy, activity and number of connection at a single screen.

3 Analytic Processes

In order to create a pixel-oriented Treemap, we pre-processed the MC1 dataset by sorting the online devices based on regions and branches. This process allows patterns to be displayed on the Treemap appropriately.

The analysis was carried out by two graduate students in computer science enrolled at Virginia Tech. Both students are authors of this paper. As shown in Figure 1, the users used one large, high-resolution display, a laptop and a desktop. Our large display is arranged in a 4 × 4 matrix of 20.1 inch flat panel monitors (4 × 4 × 1600 × 1200 = 30,720,000 pixels across LCD monitors).

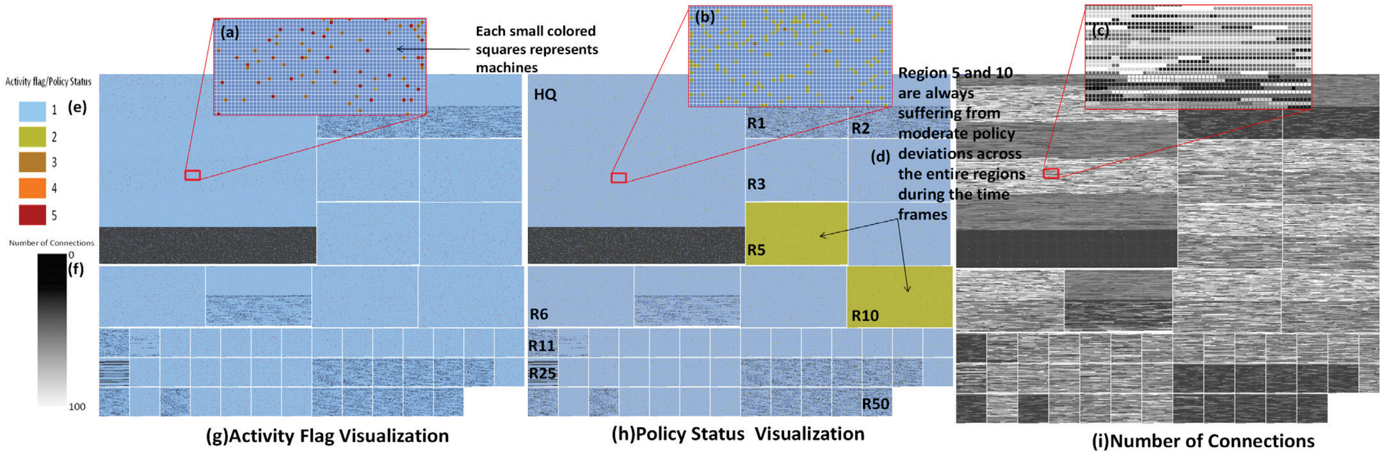

Figure 2. Three types of The Overview Treemap. Each Treemap contains statuses for 14:00 on Feb. 2. Each rectangle represents a different region in BankWorld, and small pixels (colored squares) within the rectangles present the online machines.

The two users began their analysis by examining interesting anomalies across the regions and branches using the Treemap on the large display together. For example, the users found that regions 5 and 10 (R5, 10) were constantly suffering from a moderate policy deviation in most of the machines (Figure 2d). Also, from 12:15 on 2/2, they found a sequence of computers in region 25 that were disconnected from the network. All computers in branch 33 of region 25 went offline at the same time and this anomaly spread to other computers within region 25 (R25) over time. This outage continued and the region clearly showed different patterns of offline computers from other regions by 2:15 on 2/3.

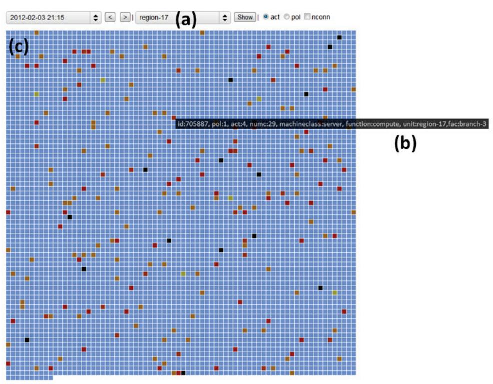

Figure 3. The Detail View of an individual machine from Region 17 at 21:15 on 2/3.

Another high-level pattern we observed on the overview of the large display is a number of invalid login attempts (orange pixels) and the addition of new devices (red) for activity flags (Figure 2a, g). For policy status, several moderate policy deviations (yellow) were also spread out across many regions and branches along with more serious policy deviations and patch failings (Figure 2b,h).

In many cases, the users used different analytic strategies using two displays. Instead of relying on a single display, our visualization facilitates utilizing some benefits of each display's visual and interaction affordances compensating the shortcomings each other.

They made some hypotheses based on basic information and suspicious changes at some regions from the Overview on the large display, and they verified more detail information for that region in the laptop and desktop computer to test and support the hypotheses. For example, the users hypothesized that the online devices would turn on during office hours and shut off at off-hours. However, from the large display, the users observed that they turned on at strange hours. Despite business hours of Bank of Money being 7am to 6pm, at 10:45am on 2/2, most of the devices were still turned off, shown in black regions (Figure 2g,h,i), but the devices were turned on all of sudden around 8pm on 2/2. So each user investigated the Detail view on the laptop and desktop to find which types of devices were turned off at this time. We learned that most of these devices were automatic teller machines and they were being used actively after finishing the normal office hours of the bank.

The users also used video to find some dynamic correlation and synchronized patterns among three different types of status over time. For instance, video allowed the users to quickly see a relationship between activity flags 3 and 5 and the number of connections (see synchronized stripe patterns between Activity flag and Number of connections in both Figure 2g and Figure 2i).

4 Discussion

Physical navigation was primarily used to explore data on the large display. Even though the Overview Treemap supports panning and zooming through mouse movement and the magnifier lens in Windows 7, all users used only physical navigation to examine the visualization. They also frequently switched between the Overview and Detail view on their laptop and desktop. After they discovered specific regions and machines to investigate through the usage of the large display, the user returned to their personal laptop and continued analysis using the Detail view.

The main limitation of our analysis was not including important domain information such as banking and geographic data. Such domain specific and geographic information in conjunction with our visualization would have provided better analysis results for the entire scenario. However, we believe that our multiple display visualization would have been able to deal with such new visual accommodations by adding more displays or new visual representations on the Treemap.

References

[1] B. Shneiderman, and M. Wattenberg, "Ordered Treemap Layouts," in Proceedings of the IEEE Information Visualization , 2001, pp. 73.

[2] R. Ball, C. North, and D. A. Bowman, "Move to improve: promoting physical navigation to increase user performance with large displays," in Proceedings of the ACM CHI, 2007, pp. 191-200.