ENTERPRISE NETWORK MONITORING USING TREEMAPS

Joseph H. Goldberg and Jonathan I. Helfman

Advanced User Interfaces
Oracle Corporation
Redwood Shores, CA

Treemaps may provide significant advantages over tabular data in corporate enterprise applications, due to their inherent ability to support users' integration of multiple dimensions of information. This study investigated the usability of treemaps for enterprise system administrators who monitor servers and web applications. Manipulated factors included hierarchy representation, data scale, and comparison with unsorted tables. The treemap hierarchy representations differed significantly in their support of the identification, comparison, and analysis tasks, but were significantly faster and more accurate than tabular data views. Treemap learnability was at least as successful as for tables. Performance differences between treemaps and tables increased with increasing size of datasets. Users' subjective ratings overwhelmingly supported treemaps over tabular data views. These results suggest that treemaps should be included as a standard graphical component in enterprise-level data analysis and monitoring applications.

INTRODUCTION

Treemaps

A treemap is a technique for visually representing data, in which the area and color of rectangles represent two different metrics (Hartigan and Kleiner, 1981; Shneiderman, 1992; Friendly, 2001). A treemap may also use nested rectangles to represent hiearachical relationships. Because of limited research, little is known about the fundamental usability of treemaps. Most treemap studies have compared algorithms for positioning and sizing rectangles (e.g., Bruls, Huizing, and vanWijk, 2000; Shneiderman and Wattenberg, 2001; Coulom, 2002; Bederson and Shneiderman, 2002). Additional studies have compared treemaps to each other (Chintalapani, et al. 2004) and to other hierarchical representations such as circular/radial trees (Stasko, et al. 2000; Kobsa, 2004), but none have focused on comparing the usability of treemaps to tables. There may be learning issues with treemaps (e.g., Stasko, et al. 2000; Chintalapani, et al. 2004), which could be a concern for adoption in domains where tables are widespread and familiar, such as database applications.

Factors Underlying Treemap Usability

The graphical representation of nested groups may influence users' comprehension of a treemap. There are two common ways to represent hierarchy in treemaps: (1) The "Classic" approach (Figure 1) uses a floating title in the center of each group, and (2) the "Titlebar" approach (Figure 2) uses separate title bars for each group. In both cases, positioning the mouse cursor over a group title highlights the group's border and provides a tooltip with the group's aggregate statistics. Multiple embedded levels of hierarchy will result in multiple floating titles or title bars, influencing both data visibility and selectability. Floating titles may obscure significant portions of underlying rectangles, and other floating titles, while the titlebar approach uses space that could otherwise show data.

The number of rectangles that are represented in a treemap may influence users' ability to identify and comprehend data trends. The present study tested two different dataset sizes: Figure 1 shows a treemap with 100 cells, and Figure 2 shows one with 1000 cells. While large treemaps have been tested in usability studies (Kobsa, 2004; Statsko, et al. 2000), no manipulation of dataset size has been reported.

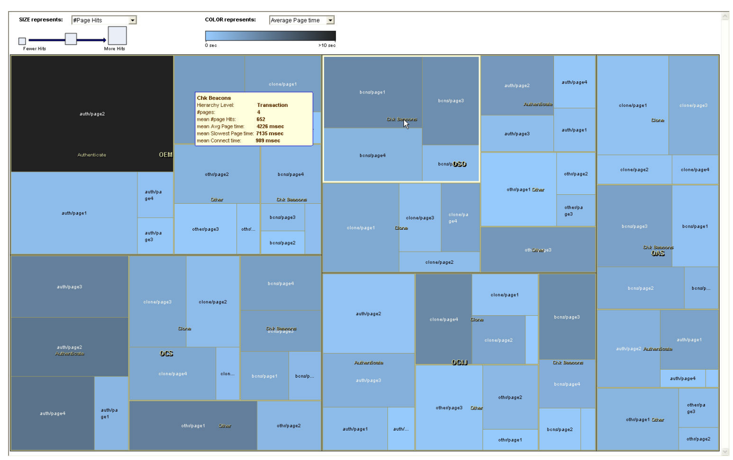

Figure 1. Classic Hierarchy Treemap, with small dataset.

Compared with tables, treemaps should provide a better high-level overview, and tighter integration between data dimensions. This advantage is expected in tasks that require some degree of rough comparison among rectangles and/or groups. When specific identification is required, especially in smaller tables, it is reasonable to expect similar performance between treemap and tabular views of the same data. Larger datasets will require time-consuming scrolling within tables, but may possibly cause more data obscuration within treemaps.

Prior studies have used treemaps to answer questions relating to the position of nodes in hierarchies, relationships among nodes and groups, and specific information identification (based upon area, color, or both). Additional factors that could influence the complexity of treemap tasks include: (1) Identification versus comparison, (2) Number of hierarchy levels traversed in comparison tasks, and (3) Whether area, color, or both dimensions are required to solve a task.

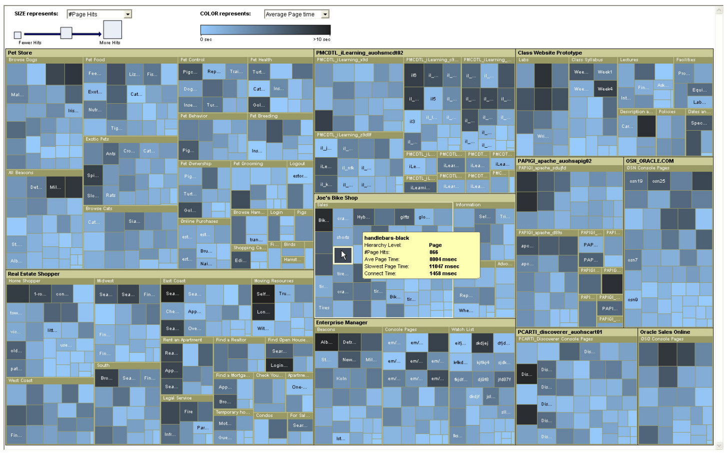

Figure 2. Titlebar Hierarchy Treemap, with large dataset.

This study investigated users' performance and subjective impressions using treemaps and tables, in identification, comparison, and analysis of hierarchically-organized data. Manipulated factors included treemap versus table data view, dataset size, and treemap hierarchy representation. Additional post-hoc factors included learnability and task complexity effects.

METHOD

Participants

Ten corporate System and Network Administrators were recruited from the San Francisco Bay Area. Each had domain knowledge of web application monitoring on a corporate network. All had some experience with graphical network monitoring tools, but none had any prior familiarity with treemaps. Each signed a disclosure agreement and was paid for participating.

The study was conducted within Oracle's Visualization Lab (Redwood Shores, CA), which has separate control and participant rooms, separated by one-way glass.

Procedure

Participants read an introductory page describing the concept of using colors and areas to convey information, the concept of hierarchy in the selected datasets, and basic performance metrics. They read a scenario statement, describing the data hierarchy and web application optimization criteria. A total of 48 tasks were then completed: 8 tasks in each of 6 conditions. Subjective rating scales were completed after each set of 8 tasks, and interviews were conducted to assess subjective impressions following the completion of all tasks.

Experimental Design and Implementation

A 3 x 2 fully-crossed factorial experiment was defined by Dataview (Treemap-Titlebar, Treemap-Classic, Table) x Dataset Size (Small, Large). Eight tasks were presented within each of these six conditions, as defined in Table 1. These ranged in difficulty from basic identification tasks, to comparison tasks, to open-ended analysis. The eight tasks were presented in the same order within each of the six conditions, but with varying questions. The six conditions (defined by data view and dataset size) were counterbalanced to assess learnability, subject to three constraints: (1) Each condition appeared an approximately equal number of times in each view order position, (2) Tables were interspersed with treemaps in view order, and (3) Large datasets were interspersed with small datasets in view order.

Table 1. Task Categories and Examples

Tasks	Task Type	Representative Examples
1, 2, 5	Identification or Counting	What was the mean Average Page Time for the Rent an Apartment transaction?
3, 4	Comparison, using one or more criteria	Which OAS page(s) had the slowest Average Page Time and the fewest #Page Hits?
7	Advanced Comparison	Which transaction contained the most pages that have more than 900 #Page Hits?
6, 8	Open-ended Analysis	Based only upon the data in front of you, what (if any) is the relationship between #Page Hits and Average Page Time?

The three dataviews included a treemap with a classic hierarchy representation (Figure 1), a treemap with a titlebar hierarchy representation (Figure 2), and an unsorted table with hierarchy represented by header indentation and color (TABL). Tables were unsorted to provide a baseline with which to compare performance on treemaps. The datasets each contained three levels of hierarchy: Web Applications (e.g., "Joe's Bike Shop"), Transactions (e.g., "Sales"), and Pages (e.g., "Pedals").

Both treemaps and tables allowed participants to freely select two performance metrics from the list: #Page Hits, Average Page Time, Slowest Page Time, and Connect Time. These were presented in two dropdowns in both tables and treemaps. In the treemaps, one was assigned to "Size" (area of rectangles), and the other to "Color" (monochromatic continuum of brightness from black to blue).

The treemaps were developed and customized from Honeycomb™ Ver. 4.8 software provided under evaluation agreement with The Hive Group (http://hivegroup.com/). The JavaScript tables were developed by the authors. All tasks were presented on a 23-inch widescreen, 16:9 ratio LCD display.

RESULTS

Completion Times

An ANOVA on completion times included terms for Dataview, Task Number, Viewing Order, and Dataset Size. Dataviews, below, are abbreviated by TM-T (Treemap-Titlebar), TM-C (Treemap-Classic), and TABL (Table). The Dataview strongly influenced task completion times, with TM-T tasks completed 30% faster (78 sec) than either TM-C (108 sec) or TABL (111 sec) views ($F=12$, $p<.001$). The increase in completion time from the small to the large dataset for the TABL view was significantly longer than the similar increase for either of the TM views (Figure 3, $F=8$, $p<.001$).

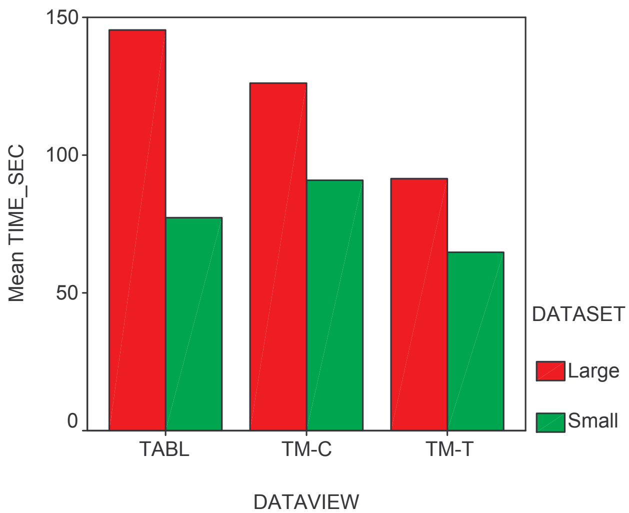

Figure 3. Task Time, by Data View and Dataset Size.

Significant differences in completion time were found between the eight tasks within a condition ($F=16.5$, $p<.001$). The TM-T format held a clear advantage within all of these task categories. Tukey's post-hoc comparisons cleanly divided the task effect into three different groups, whose completion times are shown in Figure 4: (1) Identification of values or counting of leaf nodes (Tasks 1, 2, 5), (2) Comparison across the dataset and more open-ended analysis that integrated two criteria (Tasks 3, 4, 6, 8). (3) Special comparison task ("Compare2"), requiring difficult comparison of groups, based upon lower level criteria (Task 7).

The counterbalancing scheme enabled an evaluation of the learnability of treemaps. Viewing order significantly influenced task completion times (Figure 5; $F=8.3$, $p<.001$), with the participants' first data view (153 sec) significantly slower than the second (121 sec), which was significantly slower than the remaining views (73-86 sec). Across the three data views, asymptotic performance was generally achieved by participants' third or fourth data view. Overall, the learnability of the table and the two treemaps was similar across the study, although several participants had some difficulty interpreting the TM-C in the first viewing position.

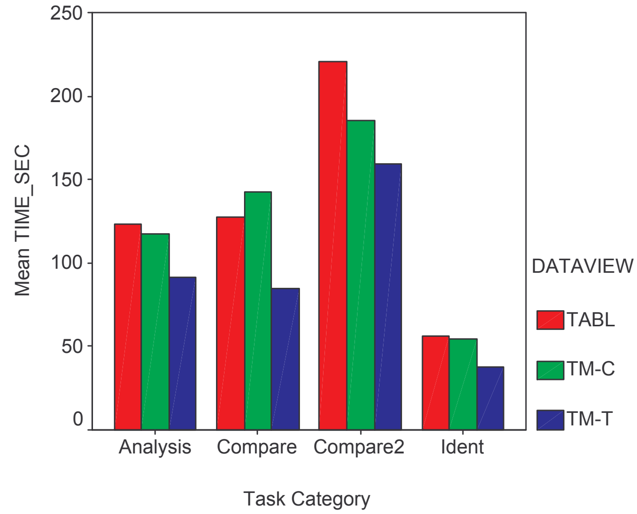

Figure 4. Task Times, by Task Category and Data View

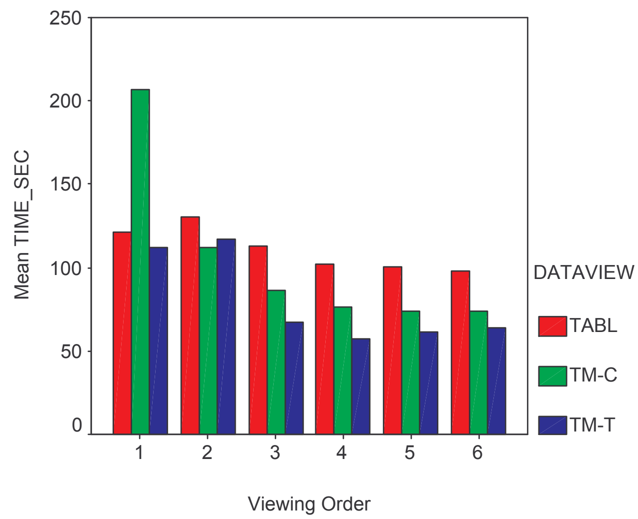

Figure 5. Task Time, by Viewing Order and Data View.

Errors and Assists

A total of 14 assists and 40 errors were recorded across 480 completed trials (10 Participants x 6 Views x 8 Tasks). Assists were provided to either: (1) Prevent the participant from performing an action that would prevent task completion, or (2) Correct misperceptions by the participant that would prevent task completion. Assists were evenly distributed across both data views and tasks, but 86% of assists occurred while participants used the larger dataset. Errors were equally distributed across the three data views, with the majority occurring in tasks 7 and 3 (57% and 25% of errors, respectively). Two-thirds of the errors in these two tasks occurred while using the larger dataset. These were also the two slowest tasks, confirming the difficulty of drawing comparisons that cross hierarchy levels.

Subjective Ratings

Following the eight tasks within each condition, participants filled out five, 7-point rating scales, relating to their experience with that condition. In each case a rating of 1 was negative (e.g. "not at all useful" or "very difficult") while a rating of 7 was positive (e.g. "very useful" or "very easy"). Means ratings, by specific question and data view, are shown in Figure 6.

The questions, "How attractive was it?" and "How easy to use was it?" produced very similar patterns, with overall ratings for the TM-T (mean 5.2) higher than for TM-C (4.6), which was in turn higher than the TABL view (3.8). When rating, "How useful was it?", respondents felt the two treemaps (5.1-5.6) were far superior to the table (3.9). The question, "How clear and understandable was it?" produced similar ratings across the three data views, from 4.5-5.0. When asked "How would you rate your ability to accomplish today's tasks?**"** the TABL and TM-C were very similar (4.4), and were moderately lower than for the TM-T (4.8).

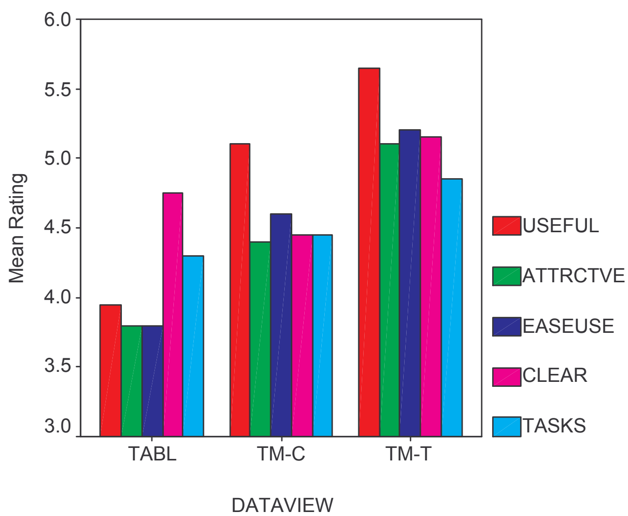

Figure 6. Mean User Ratings, by Data View, on a 1-7 Scale.

DISCUSSION

This study provided evidence that treemaps are quickly learned by system and network administrators, and are potentially superior to tabular views for identification, comparison, and analysis tasks. The titlebar hierarchy format for treemaps was strongly preferred to the classic view, which can obscure underlying information. The time and preference advantage for treemaps held for both small and large datasets.

Learnability and Task Differences

Most information visualization solutions, including treemaps, require some familiarization and learning time before users can effectively use them to solve problems. The untrained participants here all experienced significant learning in the first few data views, regardless of whether a treemap or table provided the data representation. Somewhat faster learning was evident in the mean completion times from the treemap views, than the table views. Therefore, treemaps *can* be rapidly learned by those who are completely unfamiliar with the concept, and this learning is quite rapid. The inclusion of clear legends for size and color, as well as extensive rollover tooltip information certainly added to the treemap learnability.

Four categories of tasks were embedded within the experimental tasks of this study. *Identification* tasks were completed most quickly in all three of the data views. *Comparison* and open-ended *Analysis* tasks were completed more slowly, but were similar in time to each other. Page-level *Advanced Comparison* tasks were completed the slowest in all three data views. The TM-T view was superior in completion time to both of the other views in each of these categories, whereas TM-C and TABL provided similar user performance overall. The faster TM-T performance was also reflected by users' ratings, who found it easier to use, more attractive, and more understandable than other data views.

Performance on the identification tasks was much more similar between the three data views than that from the more difficult comparison tasks. In the identification tasks, participants either scanned rows and headers in the table, or they read tooltip information in the treemaps. The major differences in task performance here were that: (1) The treemaps had data visible in one view, without any required scrolling, and (2) participants frequently had to scroll to the top of the table to identify their column metric settings.

In the comparison and analysis tasks, participants had a difficult time scanning the table view, often needing to write down values for later reference. Errors were made within this view when they failed to scroll and scan the entire table, or when they had forgotten a large value that they had previously found. The treemap views allowed easier and more accurate cross-group comprehension, and more accurate comparisons between pages and groups. However, errors were still evident because: (1) the treemap layout algorithm sized groups by the number of contained rectangles, not the metric that was mapped to rectangle area. Therefore, group areas (Web Applications and Transactions) reflected the *number* of rectangles, not the selected metric. This particularly inflated errors and assists on Task 3. (2) Participants did not always search to the very edges of the treemap, missing some important data. This was perhaps exacerbated by the widescreen views used in the study.

Data Views

It was surprising that the TABL view provided completion times nearly as fast as the TM-C view. Hierarchy levels, however, were more clearly represented in the TABL than the TM-C, where participants frequently needed to read tooltip information to determine which level they had selected. The similarity in these results emphasizes the importance of careful design in the various treemap elements. In the TM-C, floating titles that are too large can obscure underlying labels and borders. Border widths should be easily differentiable between the various levels of groups. Group and tooltip background colors should also reflect these differences in levels.

The visual and cognitive strategies used to solve the identification, comparison, and analysis tasks varied between table and treemap representations, as well as dataset size. For the treemaps, most participants preferred to search by color, rather than area. This was likely due to their familiarity with graphical "stoplight" indicators where red indicates network objects that require attention, yellow indicates warnings, and green is used for properly functioning objects. Several participants also remarked that area differences were quite subtle in many cases; precise area comparisons were no doubt difficult because most rectangles had different aspect ratios. Another reason for the color preference was the fact that area was only relevant for leaf node data, while it was sometimes possible for participants to visually judge an aggregate color for groups. One participant noted that he often tried to blur his view to see high-level trends among the groups.

The differential data densities between the large and small treemaps likely influenced participants' Useful Field of View (UFOV), or the circular area about one's fixation from which information is extracted (Mackworth, 1976). The UFOV can by estimated by the distance between two successive ocular fixations on a search task; these fixation distances decline as the visual density of background information increases. Therefore, denser treemaps likely force a smaller UFOV, and require more fixations, which in turn can lead to more missed information.

The TM-T view was superior, both by objective performance, and by subjective preference, to the TM-C view. This occurred in spite of the fact that the TM-T view reserved screen area for group names in the titlebars, leaving less space for showing leaf node data. In the TM-C view, participants did not always notice that fonts were smaller for the Transaction titles than for the Web Application title, which could have increased errors and completion times.

Recommendations

Many participants requested additional treemap features to help them in their network monitoring tasks. A user-definable threshold, beyond which leaf nodes would turn red, would allow them to more easily monitor network objects that are problematic. Automatic, system-generated refresh is required at administrator-defined intervals, such as every 30 minutes. Filtering on one or more specific metrics will enable rapid analysis of problems. Providing a historical view of a data metric is also required by some, to better diagnose problems. Some participants also remarked that a continuous color scale was not needed, in that discrete color values were sufficient for most monitoring and diagnosis tasks.

Treemaps were superior to unsorted table views in identification, comparison, and analysis tasks in this study. This advantage increased when two information metrics had to be integrated when discovering trends. Here, the titlebar representation of hierarchy was superior to a classic, floating-title representation, because of less obscuration and confusion, but specific design issues, such as fonts and colors are critical in defining these representations. While treemaps require some learning time, this was no longer than an unsorted tabular view of the same data.

Acknowledgements

The authors gratefully acknowledge the following individuals for their help in planning/developing the study, negotiating evaluation agreements, and recruiting participants: Frank Guo, Sulekha Kuthiala, Daniel Jordan, Marilyn Hollinger, Gina de Los Angeles, Anna Wichansky, Matt Thorn, Don Hoffman, Faye Kurpanek, and Tony Jewitt.

References

Bederson, B., Shneiderman, B. and Wattenberg, M. (2002), "Ordered and Quantum Treemaps: Making Effective Use of 2D Space to Display Hierarchies," ACM Transactions on Graphics 21 (4): 833-854.

Bruls, M., Huizing, K., and van Wijk, J. (2000), "Squarified Treemaps," in Data Visualization 2000, Proceedings of the Joint Eurographics and IEEE TCVG Symposium on Visualization, 33-42.

Chintalapani, G. Plaisant, C., and Shneiderman, B. (2004), "Extending the Utility of Treemaps with Flexible Hierarchy," Proceedings of International Conference on Information Visualisation, London.

Coulom, R. (2002), "Treemaps for Search-Tree Visualization," The Seventh Computer Olympiad Computer-Games Workshop Proceedings, Maasfricht.

Friendly, M. (2002), "A Brief History of the Mosaic Display," Journal of Computational & Graphical Systems 11 (1): 89-107.

Hartigan, J. A., and Kleiner, B. (1981). "Mosaics for contingency tables." in W. F. Eddy (Ed.), Computer Science and Statistics: Proceedings of the 13th Symposium on the Interface. New York: Springer-Verlag.

Kobsa, A. (2004), "User Experiments with Tree Visualization Systems," Proceedings of the IEEE Symposium on Information Visualization (INFOVIS '04), pp. 9-16.

Mackworth, N.H. (1976), "Ways of recording line of sight," in Month, R., and Senders, W. (Eds.), Eye Movements and Psychological Processing, Hillsdale, NJ: Erlbaum.

Shneiderman, B. (2004), "Treemaps for space-constrained visualization of hierarchies" http://www.cs.umd.edu/hcil/treemap-history/

Shneiderman, B. and Wattenberg, M. (2001), "Ordered Treemap Layouts," Proceedings of the IEEE Symposium on Information Visualization (INFOVIS '01).

Shneiderman, B. (1992), "Tree visualization with treemaps: a 2D space-filling approach," ACM Transactions on Graphics, 11 (1) : 92-99.

Stasko, J., Catrambone, R., Guzdial, M., and McDonald, K. (2000), "An evaluation of space-filling information visualization for depicting hierarchical structures," International Journal of Human-Computer Studies, 53, 663-694.