Log-based Longitudinal Study Finds Window Thrashing

Jock D. Mackinlay and Christiaan Royer
Palo Alto Research Center
3333 Coyote Hill Road, Palo Alto, CA 94304
{mackinlay, royer}@parc.com

ABSTRACT

Although large displays are becoming more cost effective, most user interfaces are optimized for a single monitor of modest size even though many traditional workspaces such as desks and workbenches are much larger and some studies have found benefits from large displays. This paper explores whether a single monitor is sufficient for information work using standard software. A log-based longitudinal field study finds that most of the time a single monitor allows skilled information analysts to have a reasonable pattern of window activity. However, a novel visualization of the data shows that windows typically fill the monitor and the pattern is occasionally interrupted by window thrashing, the rapid manipulation of windows caused by limited display resource. Given these findings, we identify some common tasks that justify the development and the expense of wideband visual interfaces that are optimized for larger displays.

Author Keywords

Window activity, multiple monitors, large displays, wideband visual interfaces.

ACM Classification Keywords

H5.2. Information interfaces and presentation (e.g., HCI): User Interfaces *Graphical user interfaces (GUI), Screen design (e.g. text, graphics, color), Windowing systems*.

INTRODUCTION

Sellen and O'Hara describe field and lab studies of knowledge workers engaged in reading and writing (a sensemaking activity) that found that the spatial layout of documents was used to gain a sense of the overall structure of documents, to cross-reference documents, and to interleave reading and writing [3,4]. Paper documents, in particular, were laid out in space to visualize a great deal of information, placed in juxtaposition for cross-referencing, and organized into independent reading and writing spaces that could be accessed concurrently and manipulated independently [3].

These findings suggest that workspace size can impact the effectiveness of external cognition. Although people can occasionally use airplane tray tables to work successfully with paper documents, they typically use desks and sometimes need a dining table or something larger. However, typical computer workspaces are no larger than airplane tray tables. The Sellen and O'Hara lab study found that subjects working with online documents experienced lost resolution through zooming documents, overlapping windows, and difficulties integrating reading and writing [3]. They concluded that multiple monitors might improve the reading and writing of online documents.

However, few users have multiple-monitor computers, even though such products have existed for many years. This is surprising—a field study by Grudin found that multiple monitor use confers many benefits, including peripheral awareness and improved access to resources [2]. Furthermore, Czerwinski et al review many studies that also indicate benefits for increased display size [1]. Their recent study found that a large research display had significant benefits over a standard LCD monitor for complex multi-application computer tasks. Grudin, in particular, suggested that one reason for the lack of interest in multiple monitors is that applications do not make good use of multiple monitors. Dialog boxes are placed randomly, sometimes across seams. Grudin concluded that there are opportunities for user interface designs that take advantage of multiple monitor scenarios.

This paper explores whether a single monitor is sufficient for information work using standard software. We describe a log-based longitudinal field study that found that most of the time a single monitor allows skilled information analysts to have a reasonable pattern of window activity. However, a novel visualization of the data shows that windows typically fill the monitor and the pattern of window activity is occasionally interrupted by the rapid manipulation of windows. We call this pattern *window thrashing* because it is similar to the thrashing that occurs in virtual memory operating systems when an application needs more data than fits simultaneously in main memory [5]. Given these findings, we identify some common tasks that justify the development and the expense of user interfaces that require larger displays.

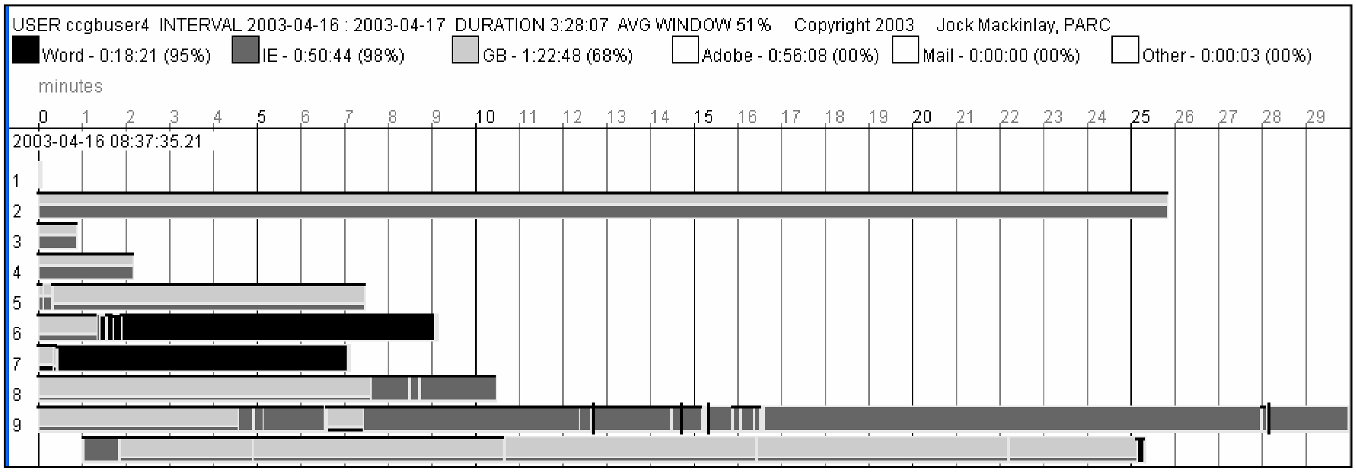

Figure 1: These 9 sessions show typical window event behavior. The horizontal axis is minutes. The vertical axis is sessions. Fill indicates window type. Bar length is segmented into window events, showing their duration. Bar height is segmented into windows, showing their display area. Rectangles with dark tops represent focused windows. Sessions 1-5 show notes written in the focused Glass Box application (GB), which covers around half of the display. MS Internet Explorer (IE) fills the rest of the display. Sessions 6 and 7 start in the GB and switch to MS Word, which covers the whole display. Session 8 and 9 start in GB and switch to IE. Session 9, which is wrapped to a second line because of its length, concludes in the GB application. Report UIR-2003-05 at www.parc.com/istl/projects/uir/publications/index.html includes a color version of the visualization. ***Analyst 4 on 4/16.***

LONGITUDINAL FIELD STUDY OF WINDOW ACTIVITY

Background

Novel Intelligence from Massive Data (NIMD) is a government research program focusing on information analysts and how to improve their work. A notable feature of this program is the Glass Box Analysis environment, custom software that instruments Microsoft Windows to capture a large amount of data including mouse events, keyboard events, window events, documents, queries, notes, and the topics the analysts are researching. A goal of the NIMD program is to use the Glass Box to capture the activity of information analysts both before and after the introduction of new technologies.

Method

The longitudinal field study of window activity described in this paper examines the baseline activities of two highly-skilled information analysts paid to do their typical work, which is using open source material obtained from the Internet to write reports. The subjects are called Analyst 3 & 4 (more analysts are planned in the NIMD program). The Glass Box environment was running on a standard PC in a typical office setting with Internet access and one 1280x1024 monitor. Human subject guidelines were followed in the development and deployment of the Glass Box. In particular, subjects can use the Glass Box application to protect their privacy by turning the capture off and on at any time, which groups the data into sessions.

The data used in the paper was captured from April 16 to July 31, 2003 and stored in an SQL database (see Table 1). The activities of Analyst 3 were captured on 33 days (more than 106 hours) and the activities of Analyst 4 were captured on 52 days (more than 269 hours). Given the volume of data captured by the Glass Box, we focus here on window events, which describe when windows are created or destroyed and when they change their size, position, or title. Window scrolling is not noted by window events but may someday be inferred from the mouse events also captured by the Glass Box.

Given the quantity of data collected, we developed custom software to filter and visualize the data. Interpreting Microsoft window events is not entirely straightforward because applications also use windows as part of their user interface widgets. A single user interaction can thus produce many confusing window events associated with nested window widgets. We removed many of these confusing events by filtering an event when the associated window had a parent window or did not have a title. For the remaining events, we determined the top level parent window and used that window for the analysis reported in this paper.

Analysis

We developed a novel visualization to analyze window manipulation activity (see Figure 1). The visualization was generated by sorting the window events into temporal order and simulating the windowing activity by interpreting events such as creation, resize, and destruction. Although sessions start with unknown window configurations, the Glass Box reports on the size and position of all windows at the start of each session, which is used to update the window simulation. The rectangles in the bars of the visualization represent windows. Their height represents the percentage of the monitor covered by that window. Their length represents the time to the next window event.

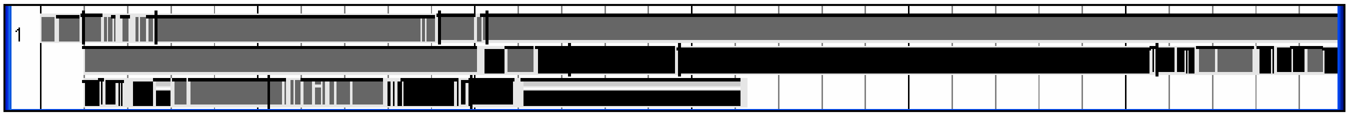

Figure 2: This session, which is also typical, started with IE window events of short duration, which is probably browsing. Next there are longer events in IE (reading) and Word (writing). Vertical red lines indicate a new window for the same type of focused application. The rapid switches between applications toward the end of this session may be window thrashing. ***Analyst 3 on 4/17.***

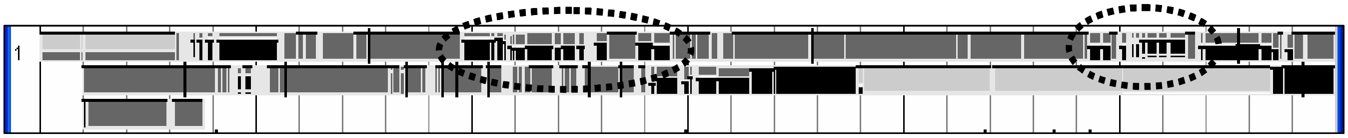

Figure 3: Window thrashing by space-multiplexing Word and IE. Word, in particular, covers only part of the monitor even though it is generally used full screen and it is resized, presumably to see information in the overlapped IE window. This visualization does not show window movements, which would be another indicator space multiplexing. ***Analyst 4 on 6/9.***

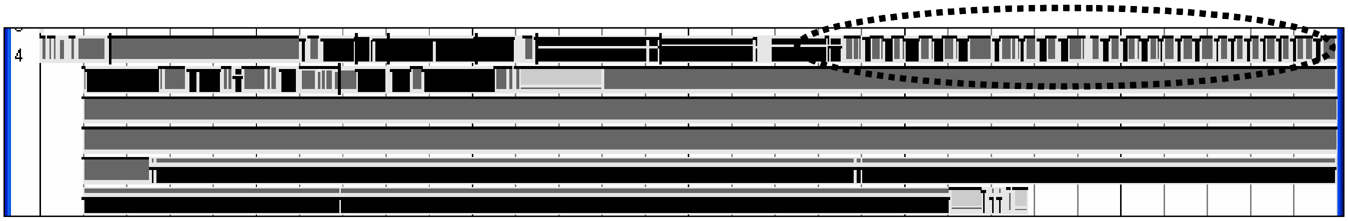

Figure 4: Window thrashing by time-multiplexing Word and IE. Time multiplexing is more common.. ***Analyst 4 on 7/1.***

The primary advantage of this novel visualization is that it compactly shows both the frequency of window events and monitor usage. Although the data is full of unexpected system-generated events and simultaneous events for a single user action, the visualization emphasizes events that are generated with a frequency consistent with human activity. Given additional evidence from samples of the mouse and keyboard events, we believe the pattern of activity in Figure 1 represents typical reading and writing. Since the window title changes when new pages are fetched into IE, we believe the short duration events at the start of Figure 2 shows browsing activity. The full activity of Analyst 3 prints on a 40 by 40 inch plot. Analyst 4 prints on a 40 by 1440 inch plot. When we look at these plots, most of the time is spent in activities that have reading, writing, and browsing patterns similar to Figure 1 and Figure 2. The visualization and the associated statistics (see Table 1) clearly show that the analysts spend most of their time reading and writing with windows that cover most of the 1280x1024 display. On average, the focused windows of Analyst 3 cover 51% of the monitor and those of Analyst 4 cover 72%. The finding of monitor usage is even larger when we focus on sensemaking applications. Word windows covered on average more than 90% of the monitor for both analysts, and IE windows covered almost 100% of the monitor. Reflecting their dual use, Glass Box windows covered 13% for Analyst 3 and 61% for Analyst 4. Analyst 4 clearly used this application to author meta-notes and routinely made the window large. Analyst 3, on the hand, only made it large on two days, and there were also some long events when it was on top and small that suggest that Analyst 3 had left for the day.

Analyst	**3**	**4**
Days	33	52
Duration	106:37:32	269:33:29
Avg. Coverage	51%	72%
Word	7:08:48 (97%)	59:13:34 (92%)
IE	48:02:14 (100%)	66:15:34 (99%)
GB	51:23:54 (13%)	31:27:20 (61%)
Abobe	0:0:28 (34%)	0:05:51 (85%)
Mail	0:0:6 (6%)	10:27:52 (85%)
Other	0:2:04 (5%)	2:03:40 (27%)

Table 1. Activity statistics for Analyst 3 & 4. The percentages indicate the amount of coverage of the monitor by a given type of application.

Evidence of Window Thrashing

The visualization also provides evidence that window thrashing occurs periodically. For example, in both Figure 3 and Figure 4 the analyst appears to be using Word and IE together, presumably writing about information found while browsing and reading. Figure 3 illustrates a "space-multiplexing" strategy where the analyst frequently changed the size of the Word window, presumably to reveal information in the underlying IE window. The visualization does not show repositioning of windows, which would be another indicator of space-multiplexing. Figure 4 illustrates the alternate "time-multiplexing" strategy, where the analyst switches rapidly between applications that cover 100% of the monitor. Time-multiplexing is more common in the data, probably because space-multiplexing requires window resizing, which is an expensive subtask. The extremely short switches shown in Figure 4 represent cut-and-paste activity. The short glances into IE with longer durations into Word after the cut-and-paste probably represent writing time, where it would have been useful to have separate reading and writing spaces on the display.

Implications

Although the frequency of window events indicate that the analysts spent most of their time reading, writing, and browsing, they used windows that filled most of the 1280x1024 monitor, and they occasionally experienced what appeared to be window thrashing. Given the limited time spent in window thrashing, users engaged in sensemaking tasks may not feel the need to invest in wideband displays. However, Table 2 shows that Analyst 4 experienced at least one episode of window thrashing on almost half of the days of the field study and both analysts experienced one day with four episodes of window thrashing. The thrashing episodes are generally similar to Figure 3 and Figure 4 but shorter in duration. Since Analyst 3 used IE much more than Word, there was less opportunity for window thrashing between those two applications.

Analyst	3	4
Total Days	33	52
Thrashing Days	9	22
Percent	27%	42%
Max episodes	4	4

Table 2. Window thrashing statistics for Analyst 3 & 4.

The visualizations shows that time-multiplexing is more common than space-multiplexing, which makes sense given the work required to resize and position windows when the space-multiplexing is used. Clearly, a second monitor would reduce this window thrashing by giving the analysts a space for reading and a space for writing. The study data does not indicate whether more than two monitors would be desirable.

Although our current field study data is limited to two subjects, the longitudinal nature of the data collected makes it possible to see window thrashing even though it occurred infrequently. Although the impact of the thrashing is not directly measured by the field study, we expect that the thrashing might have had considerable impact on the sensemaking activity as a whole because it occurred at critical moments in the sensemaking activity when the analysts were trying to write about the information they found when reading and browsing.

The information analysts used in this field study are atypical users of computer systems, engaging in complex sensemaking tasks that require weeks to complete and involving many documents. Since these subjects were quite skilled at using small displays, the thrashing found in this study probably represents a lower bound for thrashing activity. Less skilled sensemakers, such as students writing research papers, are likely to experience more frequent window thrashing when working on small displays.

CONCLUSION

We described a longitudinal field study of two information analysts using single monitor computers. The study, which found that windows typically filled the display and the analysts were occasionally interrupted by what appeared to be window thrashing, suggests the need for larger displays.

Although multiple monitor displays and larger displays are becoming more affordable, they still cost more than a single monitor display, which has been the default for most user interface designs. The longitudinal study described in this paper suggest there are three common activities that are impacted by a single monitor of moderate size: 1) multiple window tasks, such as reading and writing or programming, 2) large window tasks, such as spreadsheets and maps, and 3) multi-tasking, such as frequent interruptions involving email or personal information. We hope that users and researchers who frequently deal in such activities will be able to use the techniques described in this paper to justify the additional cost of larger displays.

ACKNOWLEDGMENTS

This research has been funded in part by contract # MDA904-03-C-0404 awarded to Stuart K. Card and Peter Pirolli from the Advanced Research and Development Activity, Novel Intelligence from Massive Data program. We thank Stuart K. Card for many discussions about window paradigms and Polle T. Zellweger for her skilled suggestions about multiple drafts.

REFERENCES

1. Czerwinski, M., Smith, G., Regan, T., Meyers, B., Robertson, G., Starkweather, G. Toward characterizing the productivity benefits of very large displays. *Proc. Interact 2003*.
2. Grudin, J. Partitioning digital worlds: Focal and peripheral awareness in multiple monitor use. *Proc. CHI 2001*, CHI Letters 3 (1) 458-465.
3. O'Hara K. and Sellen A. A comparison of reading paper and on-line documents. *Proc. CHI 1997*, ACM Press (1997), 335-342.
4. Sellen A. and Harper R. Paper as an analytic resource for the design of new technologies. *Proc. CHI 1997*, ACM Press (1997), 319-326.
5. Tanenbaum, T. *Modern Operating Systems* (2001) Prentice Hall.