Workload Analysis of 3D versus 2D User-Interfaces for Network Management

Mohsen Kahani

Faculty of Engineering
Ferdowsi University of Mashhad
Azadi Sq. Mashhad, Iran
E-mail: kahani@ferdowsi.um.ac.ir

H. W. Peter Beadle

Motorola Australian Research Center
12 Lord St., Botany
NSW, 2019, Australia
Email: pbeadle@arc.corp.mot.com

Abstract

Network management is one of the important issues in current networks. As network management is a continuous task, which still relies on human for critical decisions, the effect of the user interface on network operators is important. In this paper, we compare the overal taskload of two types of user interfaces for network management. A typical commercial traditional 2D system, Cabeltron Spectrum, is compared with the prototype 3D WWW-based Network management System (WNMS). The NASA TaskLoad indeX (TLX), with a new WWW-based user interface, was used for comparison.

1. Introduction

Networks are becoming increasingly more complex and automated. This is due in part to their increased bandwidth, the need for automated protection switching, the deployment of virtual networks based on B-ISDN and ATM technology and the increase in traffic complexity typified by multimedia communications. In turn, network management is becoming more complex and more mission critical to a larger number of organisations.

Network management systems employ software and hardware resources to maintain a desired level of service at all times and to maximise the network efficiency and productivity [3]. They collect the required data from the underlying network and process them. Although these systems are automated, so that they can handle trivial problems, they still rely on their operators for most critical decisions. As a result, the user interface has a crucial role for notifying the operators of the failure, and helping them fix the problem quickly and efficiently. The importance of the interface becomes more apparent if it is noted that the network management is a continuous task.

Realising this importance, Windows Icons Mouse Pointer (WIMP) based direct manipulation user interfaces have been developed and widely used for network management systems. To reduce the risk of human mistakes, we have been investigating the use of Virtual Reality (VR) user interface technology for network management applications. In particular, a WWW-based three-dimensional multiuser environment, called WNMS [6], has been developed for this purpose.

In this paper, we compare WNMS with one of the typical network management systems in the market, Cabletron Spectrum [4] in terms of the user interface workload over the user. The NASA TaskLoad indeX (TLX) [2] system has been used for the comparison. However, we have designed a new Web-based user interface for the TLX to make it easier to use.

Initially, both Spectrum and WNMS are briefly explained. Then we discuss the TLX, and its new user interface. Then we describe our methodology for performing the evaluation. This is followed by the evaluation results and their analysis. Finally, we conclude the paper by mentioning our major finding from this experiment.

It should be noted that the WWW-based system is a prototype while Spectrum is a well-developed system and represents one of the better user interfaces for network management in the market. That is, WNMS is one possible way of designing Web-based 3D systems and its disadvantages are not necessarily inherited from its three-dimensional display. Therefore, its merits show the merit of the architecture, and its pitfalls provide background for further investigations. Interested readers can refer to [5] for a complete comparison of the systems.

2. General Description of Systems

Cabeltron Spectrum is one of the well-known integrated network management systems in the market. Spectrum consists of two parts: SpectroSERVER and SpectroGRAPH. SpectroSERVER collects the management information from the underlying network using network management protocols, such as SNMP, and stores them in an object-oriented database. SpectroGRAPH uses the data gathered by the SpectroSERVER, and provides a two-dimensional graphical user interface, in which user can navigate the network and manage it. A typical view of this user interface is shown in Figure 1.

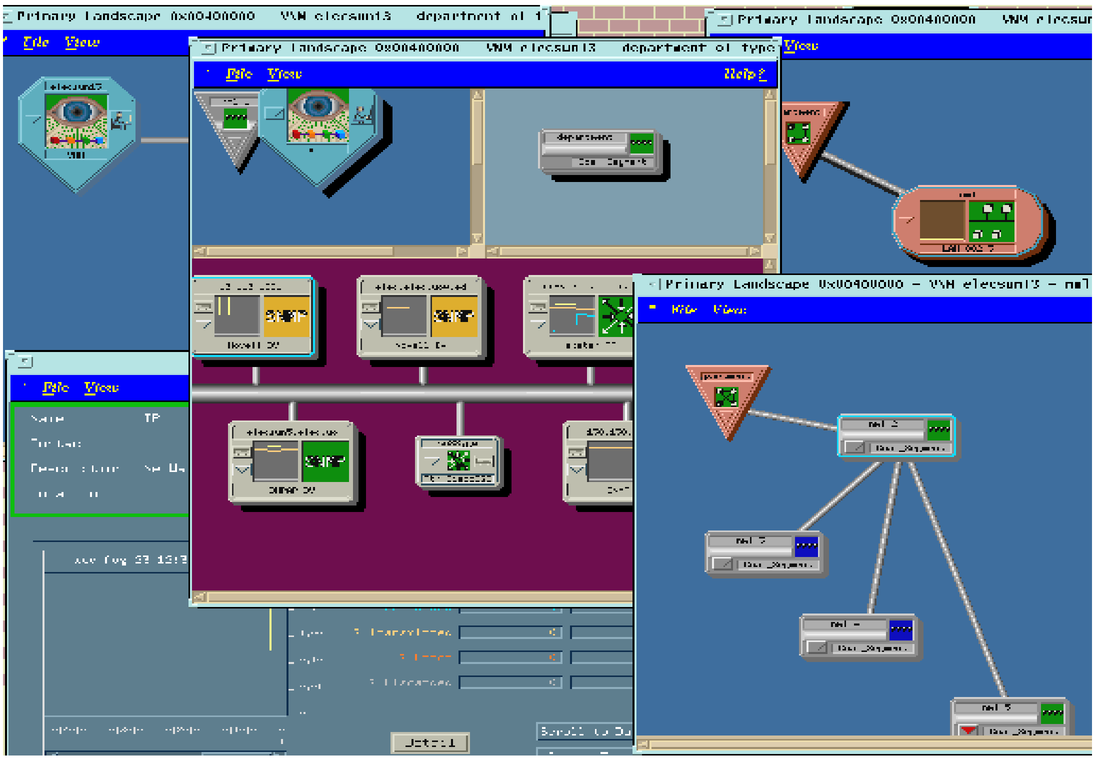

Figure 1- A typical view of Spectrum user interfaces

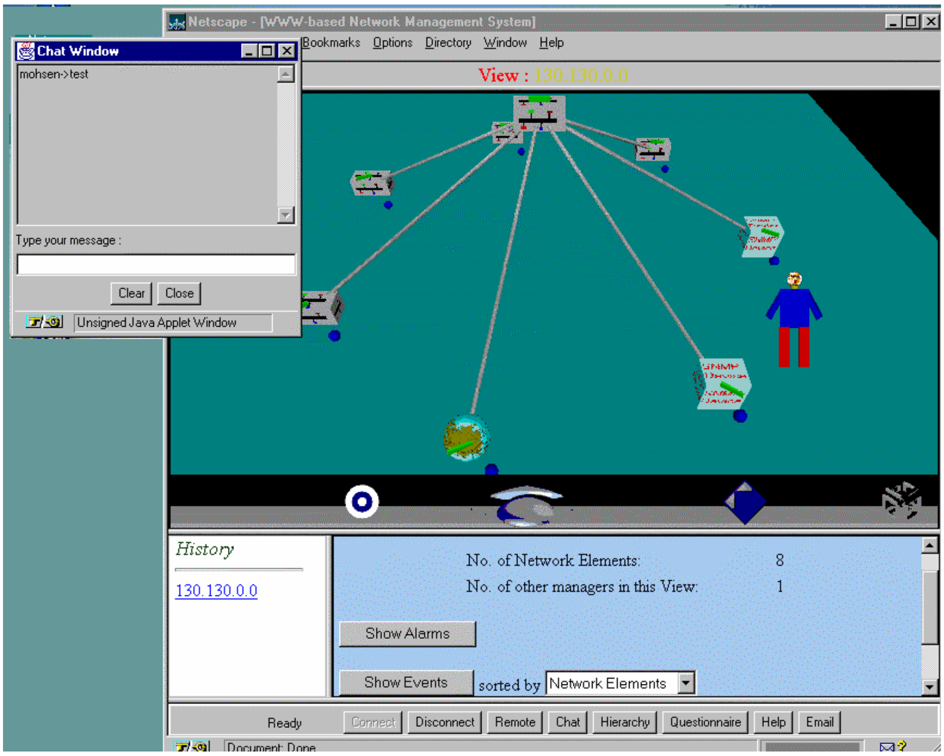

Figure 2- A typical view of WNMS

WNMS is a prototype three-dimensional virtual reality multiuser interface for management of telecommunication networks. This system can communicate with any network management system to provide a more flexible 3D user interface. It uses WWW technologies such as HyperText Markup Language (HTML), Virtual Reality Modeling Language (VRML), Common Gateway Interface (CGI) script, Java and JavaScript to build an interactive distributed multiuser interface, in which the operators potentially can manage huge networks, cooperatively. In order to make it more comparable with Spectrum, SpectroSERVER was used for collection of management data. A typical view of WNMS is shown in Figure 2.

3. NASA TaskLoad indeX (TLX)

NASA-TLX is a standard questionnaire developed by NASA Ames Research. It is a multi-dimensional rating procedure that provides an overall workload score (WWL) based on a weighted average of ratings on six subscales: Mental Demands (MD), Physical Demands (PD), Temporal Demands (TD), Effort (EF), Own Performance (OP), and Frustration (FR).

Mental demands determine the amount of required mental and perceptual activities, such as thinking, deciding, calculating, remembering, etc. That is, if the task was mentally easy or demanding, simple or complex, exacting or forgiving.

Physical demands show the amount of physical activity that is required to perform the task. That is, if the task was physically easy or demanding, slow or brisk, slack or strenuous, restful or laborious.

Temporal demands determine the time pressure that the users felt due to the rate or pace at which the tasks or task elements occurred.

The effort scale indicates how hard the users had to work (mentally and physically) to accomplish their level of performance. For performance scale users express how successful they think they have been in accomplishing the goals of the task set by the experimenter (or themselves). Finally, the frustration level indicates how insecure, discouraged, irritated, stressed and annoyed they were during the experiment.

The original questionnaire consists of a set of DOS-based programs. Initial testing revealed that the subjects did not like the interface to the questionnaire itself. To remedy this problem, a Web-based user interface, using Java and JavaScript, with the same functionality was built for the TLX. This new interface made the use of the questionnaire much easier for the subjects. Figure 3 shows a screen shot of the new interface.

The subjects should initially rate the above six items of the questionnaire, individually. Then, they should weight those items against each other. Using the user's selection, the overall workload is calculated and the form is sent to the server to be processed by a CGI script and stored in a file. After all the subjects had completed the questionnaire, the average value for each factor is calculated.

4. Performing Evaluation

The evaluation of the taskload index was performed as part of a bigger experiment. The main experiment included qualitative and quantitative evaluation of Spectrum versus WNMS in terms of network management issues and general issues. The taskload evaluation was a part of general issues comparison.

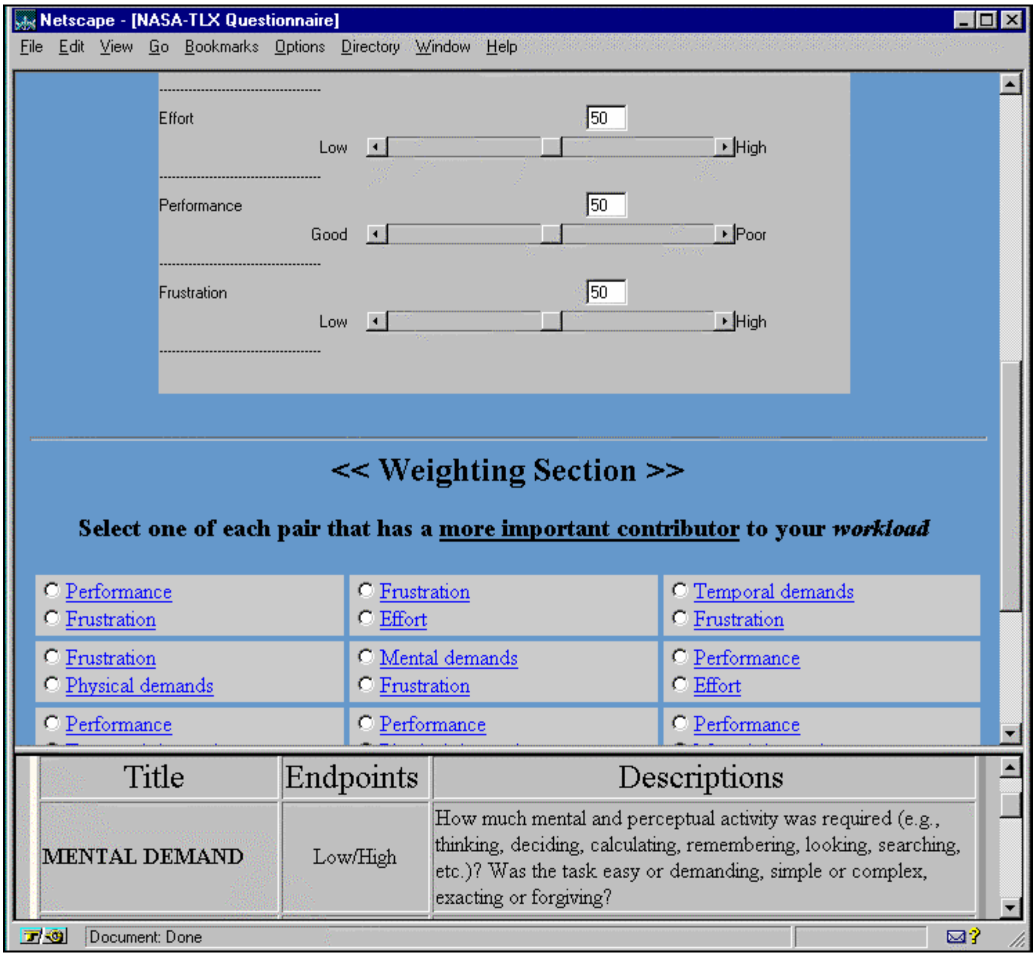

Figure 3- Web-based interface of NASA-TLX questionnaire

Overally 15 subjects were selected for the experiment. As in the main experiment several evaluation methods were performed to assess different attributes of the systems, and each attribute should have been evaluated by a set of suitably qualified subjects, the subject set was categorised into three subsets, as indicated below:

1. Novice users: The novice user set consists of five postgraduate students. While they were familiar with networking and had been using WWW for a long period, they did not have any network management background. In addition, they had no or little experience with VRML. The reason no undergraduate students were chosen was that VR and 3D user interfaces are attractive, especially for young people. This attraction could have influenced them to make a biased decision. In addition, more time would have been required to teach them how to use the systems, particularly Spectrum.
2. Intermediate users: This group consists of five postgraduate students. Although they had little network management experience, they were familiar with network management issues. Most of this group had attended network management courses, and were modestly familiar with Spectrum. Some of them also had experience with VRML.
3. Expert users: The expert set consisted of five people experienced in campus-level network management. They were familiar with Spectrum, and had some experience with VRML.

The subjects performed the evaluation individually. Each subject used each system for 30-40 minutes. The systems were evaluated in a random order. Before using any system, the subjects were given a 10-minute tutorial, and they were provided with assistance during the evaluation. Upon finishing each part, the users were asked to complete a NASA-taskload index questionnaire and interviewed. They also filled a user satisfaction questionnaire at the end of the experiment. It should be mentioned that while the sample set is small, it was difficult to find qualified subjects for a broader trial.

One-tailed paired Student-test (*t-test*) method was used for statistical analysis of the result. This test, also called the dependent-directional *t-test*, is used when the same subjects are exposed to two experiments (or before and after an experiment), with the aim of proving that one experiment has better results than the other [1]. That is, it is desirable to prove not only that the experiments have different results (reject the null hypothesis), but also a particular system is better than the other. In this paper, the *t-test* results are shown in terms of the confidence level that the *null hypothesis* is rejected. For instance, if an statement or table column is followed by $p<0.05$, it means that the probability that the statement is false is less than 5%.

Table 1- Comparison of workload indices of both systems

	WWL	**MD**	**PD**	**TD**	**EF**	**OP**	**FR**
WNMS	25	21	32	25	27	16	18
Spectrum	56	55	49	49	61	47	54
t-test prob.	$p< 0.005$	$p< 0.005$	$p > 0.05$	$p< 0.005$	$p< 0.005$	$p<0.0005$	$p< 0.001$
WNMS *better than* Spectrum	✓	✓	✗	✓	✓	✓	✓

5. Evaluation Result

The result of the experiment is briefed in Table 1. The average value for each item is shown for both Spectrum and WNMS. In addition, the results of the *t-test* statistic are shown in the table to demonstrate if the observed differences are statistically significant or not. The range of numbers is from 0-100. A small number of an item for a system indicates that users felt less pressure from that system for that item.

As the table shows, for all items WNMS has lower workload indices. In fact, in all categories, except physical demands (PD), the confidence level is quite high (a value of $p<0.005$ means that with the probability of 99.5% WNMS has a lower index). The result is exceptionally good for performance (OP) and frustration (FR), showing the satisfaction of the subjects with WNMS.

As shown in Table 1, the load index for physical demand (PD) of WNMS is not significantly better than Spectrum. This was expected as more physical activities are required to navigate and interact in 3D environments, especially when using a mouse. However, the absolute value of the index (32) is not too high to affect other indices of the overall workload. As better interaction devices emerge (eg. Phantom) it is likely that WNMS will outperform Spectrum-like WIMP interfaces in physical demands as well.

6. Conclusion

In this paper, we compared the workload indices of a two-dimensional user interface for network management with a three-dimensional one. Cabletron Spectrum, one of the leading network management software in the market was chosen as a typical WIMP-based 2D user interface system. WNMS, a prototype WWW-based network management system, was chosen as the 3D user interface.

Initially, both systems were briefly introduced. Then, NASA TLX, a standard questionnaire for workload analysis was explained. The subject selection criteria and the statistical methods used for analysis of the result were then described.

The experiment showed that WNMS provide significantly less overal workload than Spectrum ($p<0.005$). In addition, WNMS proved to have lower workload indices for mental demands, temporal demands, effort, performance and frustration subscales. The only subscale that WNMS was not significantly better than Spectrum was the physical demands. This was somehow expected as navigating a 3D environment, especially using a mouse, requires more physical activities. However, it was argued that as the absolute index value of this subscale is not too high, it does not have much effect on the overal workload index.

References

[1] Ary, D., Jacons, L., Razavieh, A., “Introduction to Research in Education”, 5th Edition, Harcout Brace College Publication, Orlando, USA, 1996.

[2] Biferno, M., “Mental Workload Measurement: Event-Related Potentials and Ratings of Workload and Fatigue“, (NASA CR 177354). Moffett Field, CA: NASA Ames Research Center, 1985.

[3] Black, U., "Network Management Standards: The OSI, SNMP and CMOL Protocols", McGraw-Hill, USA, 1993.

[4] Cabeltron Inc., “Spectrum Users’ Manual V 3.0”, 1995.

[5] Kahani, M. “The Application of Virtual Reality User Interface Paradigm for Telecommunication Network Management”, thesis dissertation, University of Wollongong, 1998.

[6] Kahani, M., Beadle, P., “WWW-based 3D Distributed, Collaborative Virtual Environment for Telecommunication Network Management”, In ATNAC’96 Conference Proceeding, Australia, December 1996, pp. 483-8.