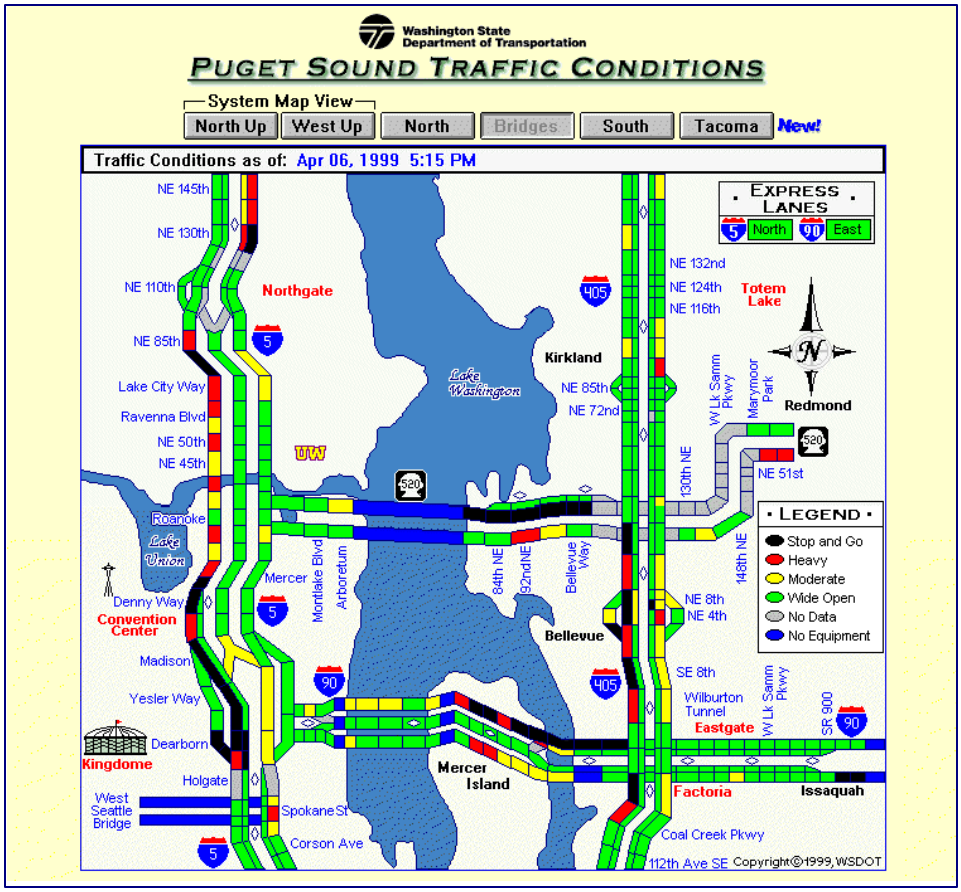
Metropolitan Model Deployment Initiative: Seattle, Washington



Travelers' Use of the WSDOT Traffic Conditions Web Site: Customer Satisfaction Evaluation

U.S. Department of Transportation
Federal Highway Administration

SmartTrek
Get Connected and Go!

January 28, 2000

1. Report No. FHWA-JPO-??-???	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Travelers' Use of the WSDOT Traffic Conditions Web Site: Customer Satisfaction Evaluation; Metropolitan Model Deployment Initiative: Seattle, Washington		**5. Report Date** January 28, 2000 **6. Performing Organization Code**
7. Authors Christopher Cluett (Battelle)		**8. Performing Organization Report No.**
9. Performing Organization Name and Address Battelle Memorial Institute 4500 Sand Point Way N.E., Suite 100 Seattle, WA 98105-3949		**10. Work Unit No. (TRAIS)** **11. Contract or Grant No.** DTFH61-96-C-00077
12. Sponsoring Agency Name and Address United States Department of Transportation ITS Joint Program Office, HVH-1 400 7th Street SW Washington, DC 20590		**13. Type of Report and Period Covered** **14. Sponsoring Agency Code** HVH-1
15. Supplementary Notes Dr. Joseph I. Peters		

16. Abstract

This report presents the results of an evaluation of public use of the Washington State Department of Transportation's (WSDOT) traffic conditions web site. This research was sponsored by the US Department of Transportation to assess customer satisfaction with the deployment of Intelligent Transportation Systems (ITS) products and services in selected metropolitan areas under the Federal Highway Administration's Metropolitan Model Deployment Initiative (MMDI) program. This report assesses traveler use of the WSDOT traveler information web site in Seattle and the greater Puget Sound area by using commercially available WebTrends™ log analysis software to analyze the web use log files that recorded individual activity on this site between May 1998 and April 1999. The URL for this web site is: http://www.wsdot.wa.gov/PugetSoundTraffic/.

The objective of this web use analysis is to better understand preferences for traveler information, the usage levels in the MMDI site areas for information provided through the Internet, the potential for reaching the traveling public via the Internet, and the patterning of use of traveler information. Customer preferences are revealed through actual usage of the web site, and these revealed preferences have been supplemented by other forms of data collection, including surveys and focus groups, conducted under the MMDI evaluation program.

This report concludes that the WSDOT traffic conditions web page is very popular with travelers in the Seattle area, and is one of the most heavily used traffic information web sites in the nation. Unusually congested traffic, caused by severe weather or incidents, prompts significant "spikes" in usage of the web site, indicating substantial latent demand for real-time traffic information. The camera image pages are particularly popular among users. The daily patterning of use clearly indicates heaviest usage during the afternoon commute peak period. Coupled with information from focus groups with active web site users and an on-line survey of users, this analysis enriches our understanding of how travelers are using the Internet to obtain up to date traffic information in their efforts to deal with some of the most severe congestion in the nation.

Key Words	18. Distribution Statement
Intelligent Transportation Systems (ITS), Metropolitan Model Deployment Initiative (MMDI), Evaluation, Customer Satisfaction, Traffic Conditions Web Site, Washington State Department of Transportation (WSDOT)	No restrictions. This document is available to the public from: The National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21.No of Pages	22. Price
Unclassified	Unclassified	25	N/A

Acknowledgements

This report was prepared under contract with the US Federal Highway Administration as a part of the Customer Satisfaction evaluation of the Metropolitan Model Deployment Initiative, under the direction of Dr. Joseph Peters, FHWA Joint Program Office. The author would like to acknowledge the contributions of Jane Lappin (EG&G, Volpe Center); Lisa D'Ambrosio (Volpe Center); Sadie Johnson (Battelle); and Peter Briglia, Greg Leege, Wayne Szydtowski, and David Richards (WSDOT).

Table of Contents

List of Figures

EXECUTIVE SUMMARY

The Washington State Department of Transportation's (WSDOT) Internet traffic conditions web site was initially established in the mid-1990's to provide travelers in the Puget Sound area with current traffic conditions and other traveler information. Traffic information web sites are being used around the country as valuable sources of information about traffic conditions, incidents, construction, traffic speeds, and links to other relevant data to help travelers make more efficient, safer, and better informed decisions. The Customer Satisfaction team is evaluating how travelers are using this information at the Metropolitan Model Deployment Initiative (MMDI) sites using data gathered in focus groups, surveys, and analysis of web use logs. This report assesses traveler use of the WSDOT traveler information web site in Seattle and the greater Puget Sound area by analyzing the web use logs that recorded individual activity on this site between May 1998 and April 1999.

The WSDOT traffic conditions web site in Washington State (http://www.wsdot.wa.gov/PugetSoundTraffic/) has evolved rapidly since its inception into one of the leading traffic information web sites in the nation. The site offers traffic conditions maps that provide real-time traffic flow and, indirectly, traffic speeds and congestion, for the major north-south freeway system, from Everett to Tacoma, and the key bridge links across Lake Washington. Traffic data are obtained from an extensive network of loop detectors located every half-mile on the major freeways. These data are processed in the WSDOT Traffic Systems Management Center (TSMC) and configured for representation on the Internet. The management of the web site takes place in Olympia, Washington at WSDOT facilities. In addition, WSDOT has installed over 200 CCTV cameras throughout the region that provide video images of current traffic conditions to the TSMC. Still video images are made available on the web site and are refreshed between every minute and a minute and a half.

As of April 1999, about 110 centerline miles of freeway were included under coverage, and this freeway coverage is anticipated to be increased modestly, by about another 20 miles, as part of WSDOT's planned deployment. MMDI funding has supported efforts to expand public access to video images on the Internet, including adding cameras, more speed and travel time information, the incorporation of probe vehicles to generate travel times for specified roadway segments, and customized e-mail alerts for users.

Using WebTrends™ log analysis software, the Customer Satisfaction evaluation team has examined the following indicators of web usage in the Puget Sound area:

- Number of user sessions, by month, from May 1998 through April 1999.
- Relationship between user sessions, page views, and hits for this web site.
- Rate of change in use over time.
- Patterns of web use by day of the week.
- Patterns of web use by time of day for weekdays and weekends.
- Number of page views for the most popular pages.
- Ratio of page views to user sessions by day.
- Duration of user sessions by day.

The objective of this web use analysis is to better understand preferences for traveler information, the usage levels in the MMDI site areas for information provided through the Internet, the potential for reaching the traveling public via the Internet, and the patterning of use of traveler information. Customer preferences are revealed through actual usage of the web site, and these revealed preferences can be supplemented by other forms of data collection, such as surveys and focus groups.

We can draw a number of tentative conclusions from this web use analysis, recognizing that web use data are complicated by measures of use that have an uncertain relationship to actual individual travel behavior in Seattle (e.g., site access from outside the region, agency maintenance uses, server glitches), and the log files are limited in their ability to distinguish individual uses of the site from automated processes (e.g., server refresh, proxy caching). The following conclusions seem plausible, notwithstanding these inherent limitations:

- The level of use has grown significantly over the period of observation, doubling on average over this one year period of observation, and can be expected to continue to increase in the future. It is reasonable to anticipate that, as the miles of roadway under surveillance increase over time with further expansions of the WSDOT program, and as new capabilities are added to the web site, usage of the site will increase. This will be reinforced by increases in computer use, Internet access, and awareness of the availability of WSDOT's traveler information web site.
- Though the underlying trend is clearly up, the WSDOT web site is characterized by month to month fluctuations of use of the site for traffic information. Two of the months, December 1998 and March 1999, showed significant jumps in usage compared to the average of their adjacent months of about 45% and 37% respectively. We have shown (Figures 3, 5 and 6) how unusual weather likely accounts for some of this increase in December, but we have not tried to analyze these "spikes" in further detail. The important lesson from these patterns of use is that latent demand for traffic information on the web can be significant; hence, to meet the needs of all users, it is important to plan to accommodate more users than visit the site on a normal daily basis. Information emerging from the analysis of the on-line WSDOT web survey suggests that users experience a slowdown in server speed and performance during periods of heavy use. In order to retain users and attract new users, server performance is a critical factor.
- Commuters are a significant component of any urban web site user population. The effect of their use is usually reflected in a pronounced peak in use during the afternoon commute period on weekdays and to a lesser degree during the morning commute period. This pattern substantiates other data that indicate greater Internet access at the office than at home, and heavier congestion with more uncertainty over a broader period of time in the afternoons. The WSDOT use data strongly support this pattern, reflecting the effects of severe congestion on commuter travel in Puget Sound.
- As is the case with other traffic web sites, travelers have a strong preference for dynamic traveler information over other static forms of information, including frequent updates on traffic flow and speeds, along with real-time camera views of traffic conditions. Eight of the top ten page views on the WSDOT site recorded in April 1999 involved camera or video-related material. Usage is spread out over more pages on this site compared with most other traffic information web sites. WSDOT includes more dynamic information pages (200)

compared with the other MMDI sites, and use patterns indicate that Seattle area travelers are accessing and value this broad array of available information.

- Recently, WSDOT users have been visiting an average of about 17 different pages during a user session. This represents a small increase in the average number of page views per user session from about 14 per user session a year earlier, perhaps reflecting the addition of more useful pages, including new camera images. Further research will be needed to better understand this apparent trend.
- Our examination of the duration of user sessions over this 12 month period also shows an increase in the mean duration of user sessions from about 24 minutes to 28 minutes. This is a longer duration compared with observations at the other MMDI sites. However, it is difficult to interpret the duration figures, since a number of external factors and assumptions can impact the length of a user session (see also Appendix B).

Traveler information web sites are becoming very popular with state departments of transportation (DOTs) around the country, and Seattle's web site is no exception. The infusion of MMDI funding in support of better and more information that can be provided through the Internet is being well received by travelers in the Seattle area. This is reflected in the rapid growth in the use of the site over the past year. WSDOT is actively developing their WSDOT web site and expanding the range of dynamic traveler information offered through the Internet. The evidence suggests that this will continue to be of significant benefit and well received.

INTRODUCTION

The Washington State Department of Transportation (WSDOT) established an Internet web site in the mid-1990s to provide travelers in Seattle and the surrounding area with current traffic conditions and other traveler information. Traffic information web sites are being used around the country as valuable sources of information about traffic conditions, incidents, construction, traffic speeds, and links to other relevant data to help travelers make more efficient, safer, and better informed decisions. The Customer Satisfaction team is evaluating how travelers are using this information at the Metropolitan Model Deployment Initiative (MMDI) sites using data gathered in several ways:

1. Qualitative data from focus groups conducted with users of these web sites.
2. On-line survey data collected directly from web users.
3. Data on access and use of these sites, recorded by the respective state DOTs in detailed web log files.

The first two approaches seek to understand travel behavior based on what travelers tell us in focus groups and surveys about their use of the web site. The third approach is being used to help us understand customer preferences for traveler information as revealed through patterns of daily web use.

This report assesses traveler use of the WSDOT traveler information web site in the Seattle area by analyzing the web use logs that recorded individual activity on this site for a full year between May 1998 and April 1999.

APPROACH

Web servers track activity on their sites by cumulating data minute by minute into log files.[1] These files contain information about each visitor to the web site, including where they are connecting from, what time the visit took place, which pages were viewed, how long the visitor remained at the site, and which browser (and browser version) was used. The files can be very large, and the WSDOT log files are consistently over 10 gigabytes a month, which is enough data to fill more than 20 CDs. For the purposes of the Customer Satisfaction evaluation, we are using commercial software called WebTrends™ to help analyze these large log files.

The primary unit of analysis for our purposes is the User Session. Ideally, we would like to be able to assess web site use by individuals. Since we cannot uniquely identify all individual users of these web sites, a user session offers the best approximation to a single user. We know that many individuals visit the traffic web site several times a day as well as multiple times each week. Over 75 percent of the respondents to a survey of users of the WSDOT traffic information

[1] Some web sites cumulate their log data in separate files month by month, while other sites record their use data for longer or shorter intervals. WSDOT records monthly log records of web site activity, and for this analysis we used these monthly log files, along with selected analyses day by day. See Appendix A for more details on these procedures.

web pages said they had consulted this site more than 10 times in the past month.[2] Thus, user sessions can help us understand the activities of individuals at a site, including the amount and patterns of use, but we won't know the number of unique individuals who visit the site or how use of the site varies by individual characteristics.

A single user session is usually associated with a larger number of page views and hits. This is clearly illustrated in Figure 1 below. The activities of a single user at the web site, plus the automated effects of image refresh by the server and the number of images located on the pages visited determine how many page hits will be associated with each page viewed by a user during a user session. Figure 1 shows this relationship for the WSDOT web site over the month of April 1999.

The key definitions of hits, page views, and user sessions follow:

> Hit: A hit is a measure of activity on a web page that is incremented each time the page is opened and each time any imbedded image on that page is opened or refreshed. Hits reflect all individual actions taken by a user to access any portion of the web site, *plus* all activities that are instigated by the server automatically. For example, every graphic image on a page generates a hit when that page is opened. In addition, the server will "push" new data out to a page to refresh the information periodically, and each time this happens, a new set of hits will be generated. Simply keeping a page open for a period of time can result in many automatically generated hits that are unrelated to discrete actions taken by the user. One page view on a complex traffic information web page may be equivalent to 30 or more hits

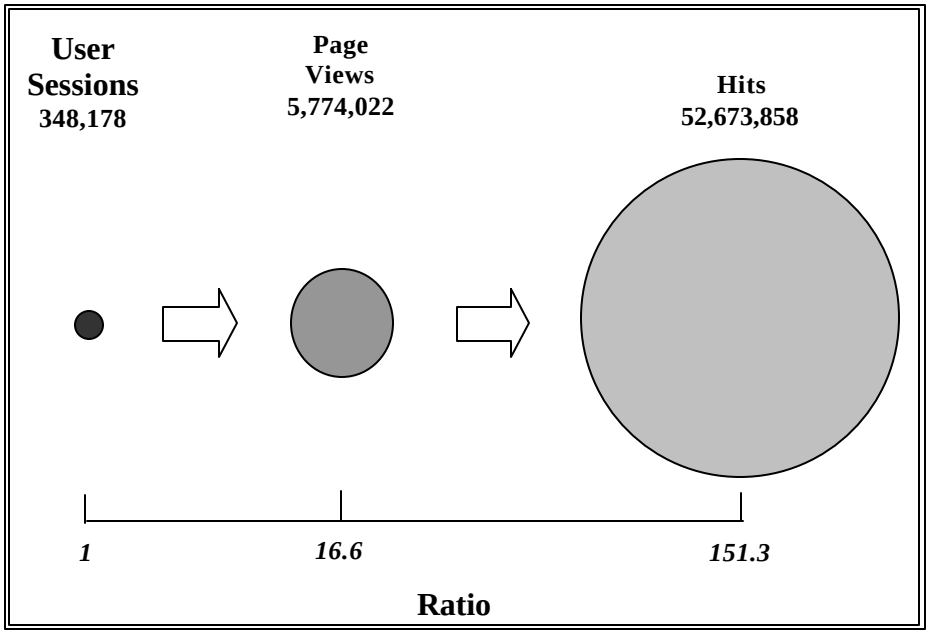


FIGURE 1. RELATIVE NUMBER OF USER SESSIONS, PAGE VIEWS, AND HITS ON THE WSDOT TRAFFIC CONDITIONS WEB SITE, APRIL 1999

[2] This finding is based on 608 respondents to an on-line survey questionnaire conducted as part of the MMDI

when that page is accessed only once. While hits are a commonly used measure of the overall level of web site activity, they turn out to be a poor indicator of how individuals actually are using the site to acquire information.

Page View: A page view is equivalent to a hit to a hyper text markup language (html) page. Since hits include activities within a page, they are substantially more numerous than the number of page views. Views give a better understanding of the kind and amount of information accessed by a user than do hits, though page views themselves also can be "pushed" by the server in order to update real time information. Thus page views, while much less numerous compared to hits, are not a perfect reflection of individual activity at a web site.

User Session: A user session reflects activity (all hits and page views combined for a period of time) for one user of a web site. A unique user is determined by the Internet protocol (IP) address. A user session is terminated when a user is inactive for a period of time that can be set for the analysis of usage patterns. For the purposes of this analysis, the 30 minute default setting was used.[3] The number of user sessions in a month is an approximation of the number of separate times (sessions) users visit a web site during that month. It is not the number of unique individuals visiting the site, as an individual can engage in many user sessions a day. A user session at a web site can vary tremendously in length of time and breadth of activity. It is in many ways the best indicator of user activity at a web site.

Depending on how many pages of different kinds of traveler information are available at a traffic web site, individual user sessions can involve visits to many different pages. While we might hypothesize that the longer a user session lasts, the more page views are likely to be accessed, this conclusion is not supported by the data in Seattle. User session duration and number of page views accessed are only weakly positively related, suggesting that other factors than how long one spends in a session are more important in explaining the number of pages viewed.[4]

Page views also are useful indicators of activity, since they represent a page of information in which the individual is interested. But just as with hits, views also can be generated automatically. For these reasons, these data need to be interpreted with caution. The data are useful for comparative analyses of web use activity over time at a site; however, using web log data to compare different sites is more problematic, given very different characteristics of web sites, their content, the local environment, and the user population.

The information recorded into web log files does not always reflect 100 percent of user activity at that web site. In some locations, end users such as television stations can access closed circuit television (CCTV) images directly without ever generating an entry in the web log, which registers only html and similar non-image pages. This process of "image capture" means that literally thousands of people can view the traffic image on their TV while no record is left to

---

Customer Satisfaction evaluation during May and June, 1999.

[3] See Appendix B for a discussion of why this default value was selected and its implications for the analysis.

[4] The correlation coefficient for three different months was calculated between number of pages viewed and duration of user sessions per day for the month. Results are 0.19 for May 1998, 0.06 for December 1998, and 0.07 for April 1999.

indicate access to that web site information. Another way that traffic web site information can be distributed to thousands of users is through a process called "proxy caching." A large company, for example, can connect to the web site and access a batch of traffic information, log off the site, and then distribute that information on their company network to all employees to view on their individual computers. The web site records one user session; yet, hundreds of people can be viewing the information. Companies do this in order to efficiently manage their limited bandwidth, but the result is that the DOT's web logs underestimate total usage by the amount of both image capture and proxy caching. This can have a significant impact on our understanding of usage levels for traffic information, and these two processes are occurring to a significant degree in Seattle during the period of our analysis.[5]

One other important aspect of web log data relates to the difficulty of detecting the origin of users. Ideally, we want to assess the use of the web site by people currently in the area or who plan to be in the area for their travel purposes. But we can only partially determine where people are when they access these web sites. Users with service providers such as America Online (AOL) are all recorded in the web log as coming from AOL, which has its headquarters in Reston, Virginia. Some users visit web sites from all over the country or from other countries. Thus we cannot distinguish users who are in or near Seattle with any accuracy.[6] The web logs give us a broad picture of levels of use of these traffic information sites and patterns of use over time. While recognizing the limitations and constraints inherent in these log data, we can gain a better understanding of awareness, interest, use, and behavior from this vast data base.

WSDOT SYSTEM

The WSDOT traffic conditions web site in Washington State (http://www.wsdot.wa.gov/PugetSoundTraffic/) has evolved rapidly since its inception into one of the leading traffic information web sites in the nation. The site offers traffic conditions maps that provide real-time traffic flow and, indirectly, traffic speeds and congestion, for the major north-south freeway system, from Everett to Tacoma, and the key bridge links across Lake Washington. Traffic data are obtained from an extensive network of loop detectors located every half-mile on the major freeways. These data are processed in the WSDOT Traffic Systems Management Center and configured for representation on the Internet. The management of the web site takes place in Olympia, Washington at WSDOT facilities. In addition, WSDOT has installed over 200 CCTV cameras throughout the region that provide video images of current traffic conditions to the TSMC. Still video images are made available on the web site and are refreshed between every minute and a minute and a half.

Figure 2 shows a snapshot image of the Freeway Conditions Map page on the WSDOT web site. The Freeway Conditions map shows the portions of the Seattle freeway network that are currently providing web site viewers with traffic information. As of April 1999, about 110 centerline miles of freeway were included under coverage, and the coverage is anticipated to be increased modestly, by about another 20 miles, as part of WSDOT's planned deployment. The

[5] WSDOT has estimated that image capture may account for as much as 30% of total web site usage.

[6] A procedure in WebTrends™, called reverse DNS lookup, offers some insight into who is visiting the web site. However, the extremely large size of the WSDOT log files renders this complex, computer memory-consuming process prohibitive.

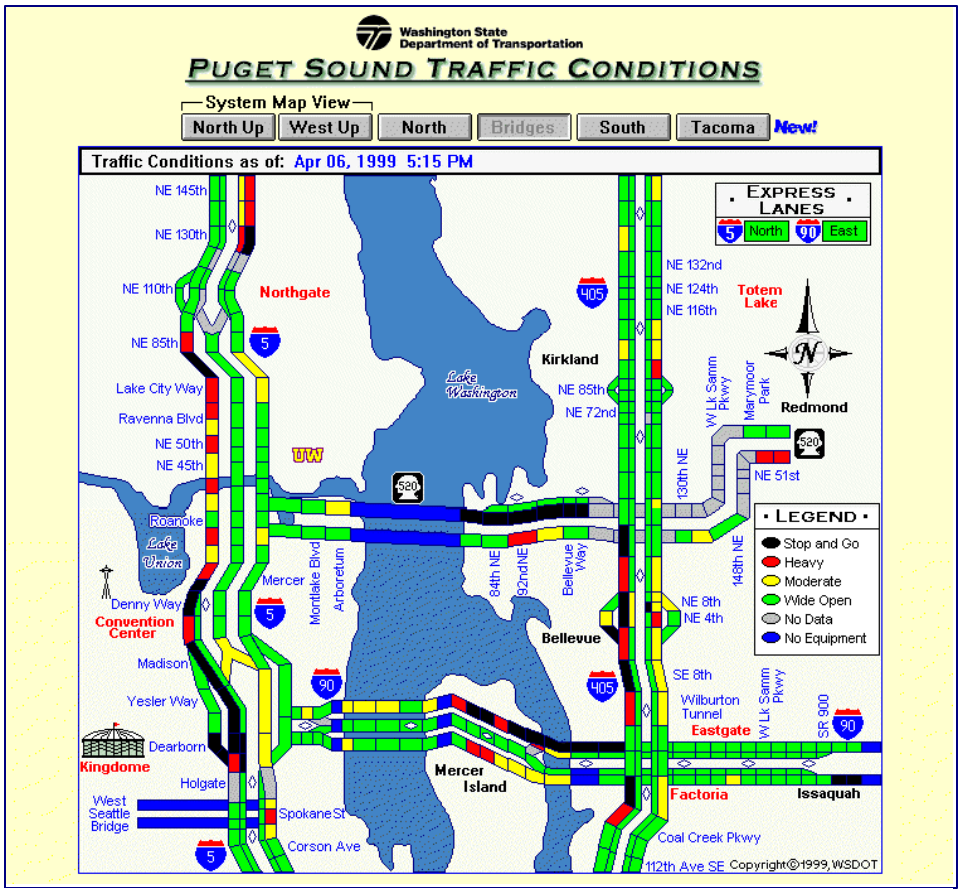


FIGURE 2. SNAPSHOT OF WSDOT WEB SITE SHOWING CURRENT PUGET SOUND TRAFFIC CONDITIONS (BRIDGES VIEW)

Traffic Conditions map (Figure 2) shows a network of automated, smart highways, developed in part with support from the Metropolitan Model Deployment Program (MMDI). MMDI funding has supported efforts to expand public access to video images on the Internet, including adding cameras on the Tacoma Narrows bridge, and arterial traffic information (SR-599). Additional planned expansions of this system will include a camera on the First Avenue South bridge, more speed and travel time information, the incorporation of probe vehicles to generate travel times for specified roadway segments, and customized e-mail alerts for users.

The WSDOT web site contains an estimated 5,000 individual pages available on the traffic information pages supported by WSDOT.[7] About 100 of these contain dynamic information. There are 25 other traffic information pages, and about 4,875 non-traffic pages (reports, pdf files, and other documents that users can read on line or print out in hard copy).

WebTrends™ allows for an analysis of how user activity at the site is linked to specific individual pages. Later in this report we will discuss which pages are most visited, suggesting which information is most popular with users of this site. For now we can point out that 87 percent of the top 50 pages viewed on this site display dynamic information on traffic conditions

[7] The number of pages reported here is based both on a links analysis by WebTrends™ and conversations with the WSDOT web masters, Mr. Wayne Szydtowski and Mr. David Richards.

in the area. This provides an indirect but compelling measure of the value that travelers place on access to real-time traffic data.

Various characteristics of a region, its population, and the available traffic information are likely to affect the levels of use of traffic information systems of any kind. The following are some of the factors likely related to levels of use of the WSDOT web site:

- The level of congestion experienced by travelers in the area. The Texas Transportation Institute (TTI) has published a congestion index for many urban areas in the U.S. Seattle's index in 1996 was 1.27, making it the sixth most congested city in this ranking.[8] TTI modified their index in their most recent report, based on 1997 data, and they refer to the new Travel Rate Index (TRI). The TRI "is a measure of the amount of extra time it takes to travel during the peak period." The TRI for Seattle in 1997 was 1.43, making the Seattle-Everett metropolitan area the second most congested among all urban areas analyzed. This TRI indicates that it takes Seattle travelers 43% longer to travel during the peak periods compared to free flow conditions.[9]
- The amount of roadway under surveillance, or the extent of traffic conditions information coverage in the area, coupled with the length of time since the system was installed. Seattle currently has an estimated 110 miles of freeway covered under WSDOT's program to provide travelers information over their web site. The web site has been in place since about 1995.
- The amount of information available on the traffic information web site. The WSDOT web site currently displays 5,000 pages of traveler information. Dynamic, real-time information is primarily provided on 200 of these pages.
- The design and functionality of the traffic information web site, including the range of information available and user perception of the accessibility, usefulness, and reliability of this information.
- The level of awareness of the population of the availability of traffic information over the Internet. This is closely connected to the amount of promotion or advertising of the site that has taken place.
- The level of access of the traveling public, both at home and at the office, to the Internet, and their degree of comfort and experience with computers and information technologies.[10]

[8] A description of TTI's Roadway Congestion Index can be found at: http://mobility.tamu.edu/study/rci.stm. This discussion states in part, "The resulting ratio indicates an undesirable level of areawide congestion if a value greater than or equal to 1.0 is obtained."

[9] Schrank, David and Tim Lomax. 1999. *The 1999 Annual Mobility Report: Information for Urban America.* Texas Transportation Institute.

[10] The 1997 "ITS Supplement" to the Puget Sound Regional Council's (PSRC) household panel survey indicates that 43% of Seattle residents have access to the Internet at home or work. However, only about 0.5% access pre-commute traffic information from the Internet at home, and about 2% say they do so at work. This survey is currently being updated. Also, the State of Washington's Office of Financial Management conducted the Washington State Population Survey in the spring of 1998 to provide social, demographic, and economic information about Washington residents. Telephone interviews were conducted with 7,279 households that represent the state population as a whole. This survey indicates that 63.0% of all households in King County report having a personal computer at home, and 47.4% report using computers to communicate (see: http://www.ofm.wa.gov/sps/findings.htm).

- The number of travelers in the population base that travel on the roadways covered by the traveler information. This is related to total population size, population density, number employed, and other demographic characteristics of the area.[11]
- The nature and extent of incidents, weather-related events, or other disruptions of traffic that would create interest in the web site.
- Characteristics of the traveler and their typical travel routes and experiences that together motivate the individual to seek out travel information. These factors may include such things as having viable route options in their daily commutes, past experience with unexpected delays and congestion (as opposed to "normal" or expected congestion), and such individual attributes as a need to be in control or a high value of time (don't like being late).
- Finally, there are likely to be errors in web log data that are hard to identify, such as brief server outages or interrupted Internet connections. Obviously, these errors increase the difficulty of accurately interpreting the data.

Next we will take a closer look at the levels and patterns of use of the WSDOT traffic information web site.

WSDOT WEB SITE ANALYSIS

Levels of Usage. For the period May 1998 to April 1999,[12] the level of usage of the Seattle WSDOT traffic information web pages ranged from a low of 236,653 user sessions in August 1998 to a high of 497,192 user sessions in December 1998. Figure 3 shows usage levels for the twelve month period as measured by the number of recorded user sessions for each month. While there has been variability month by month, in part due to major weather events that are discussed below, the general trend in usage levels has been up over the course of this period. This variability can be smoothed out by running a linear trend line through the average daily number of user sessions for the period, as shown in Figure 4. The resulting average increase in use is about 106 percent[13] per year over this time frame, or a little more than a doubling of the usage in a year. The trend line showing the average growth over the twelve month period is drawn in. Note, however, that this rate of change for this arbitrarily selected period of time may not extend into the future.

[11] The US Bureau of the Census released on June 30, 1999 the most recent population estimates for selected places (see: http://www.census.gov/Press-Release/www/1999/cb99-128.html). For cities with populations of 1 million or more, the two fastest growing cities are Phoenix and San Antonio, with population growth of 21.3% and 14.1% respectively between 1990 and 1998. Seattle grew 4% during this period to a total population in 1998 of 536,978. The 1998 population of Phoenix was 1,198,064, and San Antonio was 1,114,130.

[12] This period of analysis of usage patterns was selected to include the bulk of the MMDI project deployments, with May 1998 generally preceding those deployments and April 1999 being the most recently available month of log data.

[13] Using the average daily number of user sessions per month controls for the different lengths of months. The average rate of change per year is computed based on a linear regression line run through the 12 data points.

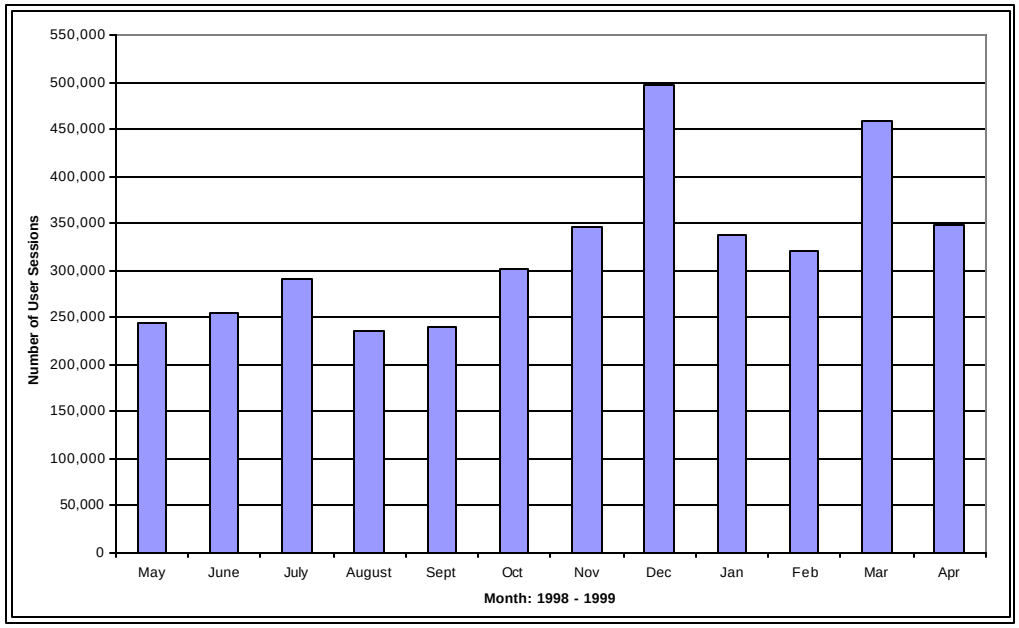


FIGURE 3. TOTAL NUMBER OF USER SESSIONS
MAY 1998 THROUGH APRIL 1999

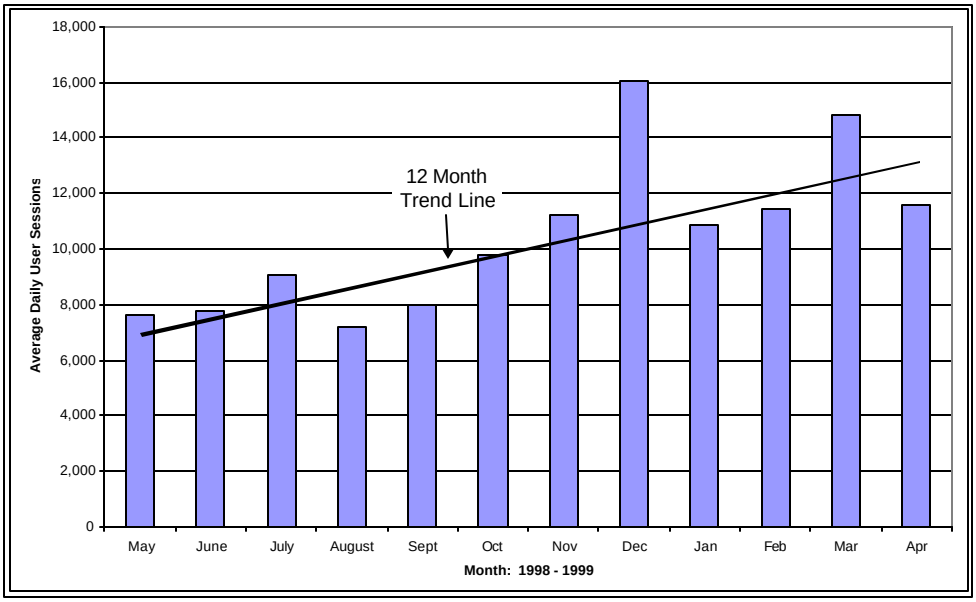


FIGURE 4. AVERAGE DAILY NUMBER OF USER SESSIONS BY MONTH,
MAY 1998 THROUGH APRIL 1999

Use Impacts of Unusual Events. Figures 5 shows the number of user sessions for each day in the month of December 1998. The average daily number of user sessions for the 31 days in this

month is 16,038, which is the highest level reached between May 1998 and April 1999, as shown in Figure 4. One of the reasons that the number of user sessions increased so much in December is because of the effects of three severe snow and ice days in Seattle. Travel was significantly restricted, there were many snow-related highway incidents, and all this occurred during the Christmas holiday. The worst impacts occurred between December 19 and 21, and Figure 5 reflects the component of the traffic web site activity that can reasonably be attributed to this event. These effects are all the more noteworthy given that two of the three days were weekend days, a period of much lower web usage. To judge the magnitude of this effect, usage during these three days was compared with the usage levels that would have occurred if web site access was the same as the two comparable periods earlier that same month. Using this assumption, the "spike" in usage attributable to the snow and ice effect is estimated to be about 57 percent, or more than half, of the total usage for those three days, as shown in Figure 5. This accounts for about 9 percent of the total number of user sessions recorded for the month of December, as shown in Figure 4. Similar winter driving weather on other days and the complexity of holiday activities clearly resulted in higher than "normal" usage levels in this month; however, we have not attempted to disentangle all these effects.

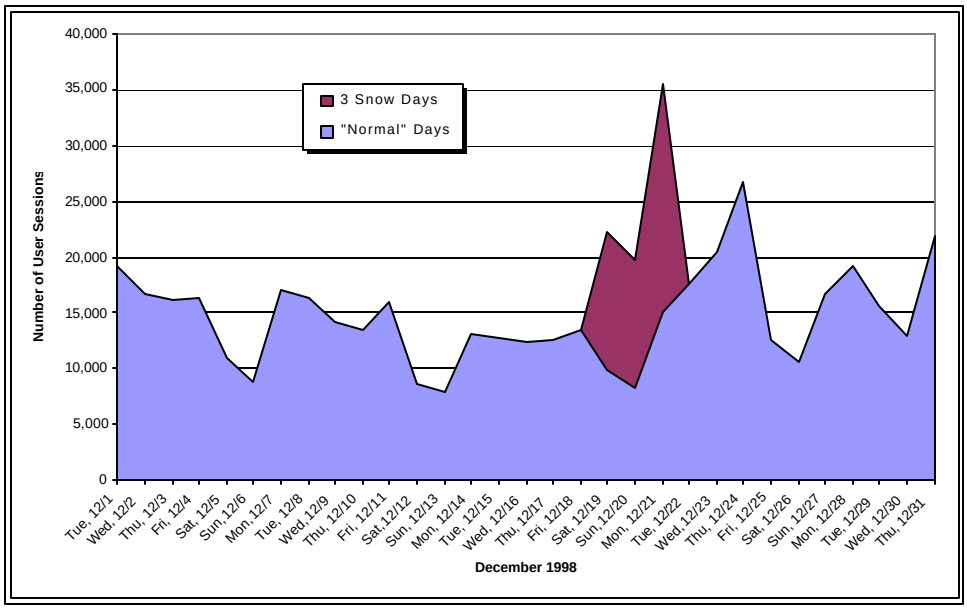


FIGURE 5. NUMBER OF USER SESSIONS PER DAY, DECEMBER 1998
THREE SNOW AND ICE DAYS IN SEATTLE

We have examined another measure of web site usage in Figure 6 for the same month of December. Here we see an even more dramatic spike in the number of pages viewed during these three days. About 71 percent of the total page views on the three snow days can be attributed to the weather. From Figures 5 and 6 it is clear that page views increased more than the number of user sessions; that is, the evidence suggests that, while more users were coming to the site than normal to obtain traffic information during this snow event, the more pronounced effect was the greater number of pages of information accessed by those users. The ratio of page views to user sessions was almost double the average of 15.2 pages viewed per user session for the month on those unusual days (see Figure 9 later in this report).

The December peak use day (Monday, December 21, 1998) recorded 35,541 user sessions, which represents a little more than two times the normal average *weekday* usage of 17,045 user sessions per day in December. This offers an indication of the potential pool of web site traffic information users in the Puget Sound area at this time. The web site use levels dropped down the next day to more normal levels. Thus, it would appear that there is a cadre of regular repeat users of the web site, a fact that is confirmed by results of the on-line survey of users conducted under this evaluation program, along with an equal number of occasional users who obviously are aware of the web site but are only induced to use it under atypical driving conditions.

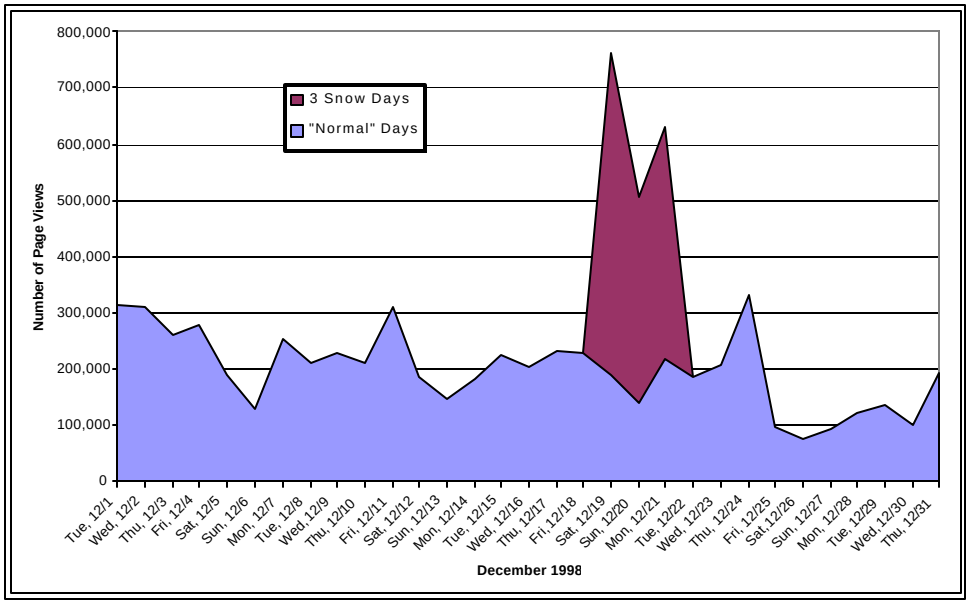


FIGURE 6. NUMBER OF PAGE VIEWS PER DAY, DECEMBER 1998
THREE SNOW AND ICE DAYS IN SEATTLE

We might hypothesize that a portion of new users who are drawn to the web site as a source of traffic information by an unusual event such as this would stay on as more regular users. The limited data observed in this example, and in the case of major weather events analyzed as part of the San Antonio MMDI evaluation, do not appear to support such an hypothesis. In each instance of a major but brief spike in use levels due to severe weather as experienced in Seattle in December and in San Antonio in October and December, levels of use fairly rapidly dropped back to the pre-event baseline levels.

There are several related questions that can be explored with the web site use data. Do unusual events tend to attract new users to the site, or do the regular users use the site differently, engaging in more user sessions per day or perhaps accessing a wider range of traffic information during their user sessions? Do the user sessions exhibit different time-of-day patterns of access during these unusual events, compared to normal days? The relatively crude nature of the web log data don't allow us to answer these questions definitively, but they are suggestive of

underlying patterns of use. In the next sections, we will look at time-of-day patterns for commuting and non-commuting periods, day-of-week use patterns, and duration-of-use patterns.

Patterns of Web Site Use. Figure 7 shows the pattern in number of user sessions by day for the months of May 1998 and February 1999. The patterns shown in Figure 7 are quite typical of the use of traffic information web sites. Use is higher during the weekdays and much lower on the weekends. Sundays are typically the lowest use days, but not always. Also characteristic of weekday use is a peak effect on Mondays and Fridays and a dip during the mid-week period, but as is shown in Figure 7, this too is not always true. The higher weekday use of the web site reflects the higher traffic levels on weekdays associated with commuting, shopping, and other normal driving patterns in urban areas.

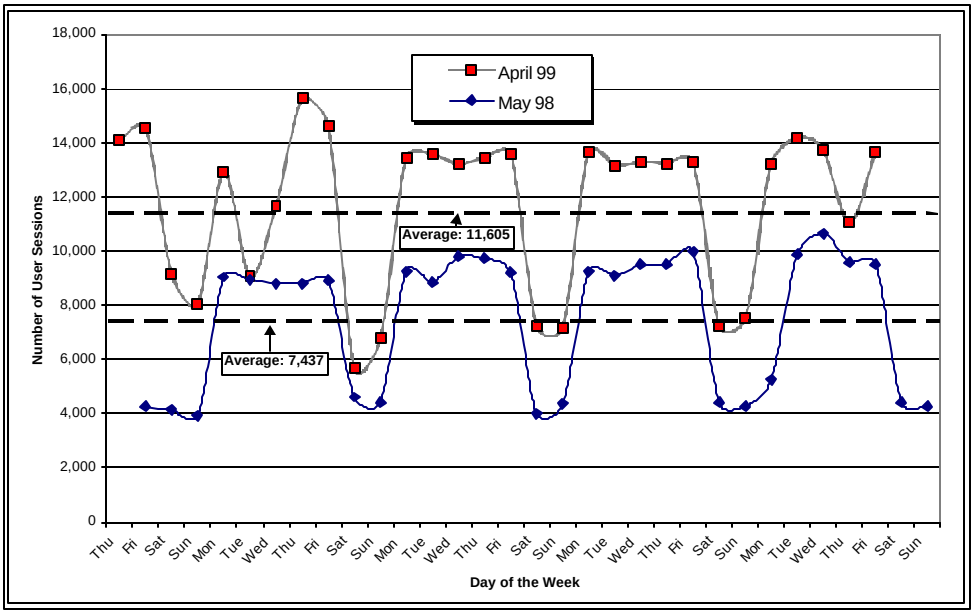


FIGURE 7. AVERAGE DAILY NUMBER OF USER SESSIONS BY MONTH, MAY 1998 AND APRIL 1999

The average daily number of user sessions that were shown in Figure 4 are also evident in Figure 7. The data are displayed by day of the week during these two months. Except for the absolute increase in number of user sessions between May and April, the daily fluctuations match closely between these two periods at either end of the 12 month period of observation. Weekend use dips below the average, and weekday use rises above the average in each of the weekly segments shown here, except for some odd fluctuation in the first week of April 1999.[14] A similar examination of usage data for several months in between May and April suggests that these patterns are consistent throughout the period, and they replicate the patterns observed at the other two MMDI sites examined.

[14] A careful examination of the hourly usage by day during April 1999 shows at least three periods with zero or very low recorded use levels, suggesting problems with the web server recording activity during these times. The days where these problems were most pronounced are those that most obviously deviate from the expected pattern in Figure 7.

Figure 8 explores further the patterns associated with use of the WSDOT traffic web site. For May 1998 and April 1999, we can observe how use of the web site varies by time of day for both weekdays and weekends, and how this pattern appears to be changing over the 12 month period of observation. Figure 4 showed an average of 7,623 user sessions per day during May 1998 and an average of 11,605 user sessions per day during April 1999, an increase of 52 percent over this period of time. The data displayed in Figure 8 show the average daily number of user sessions that occurred in May 1998 and in April 1999 split into both weekday and weekend components.

What is most striking in this figure are the time-of-day use pattern differences between weekday and weekend web use activity. In cities where commuters face consistently higher congestion during the morning and evening commutes, we usually see a corresponding increase in usage of

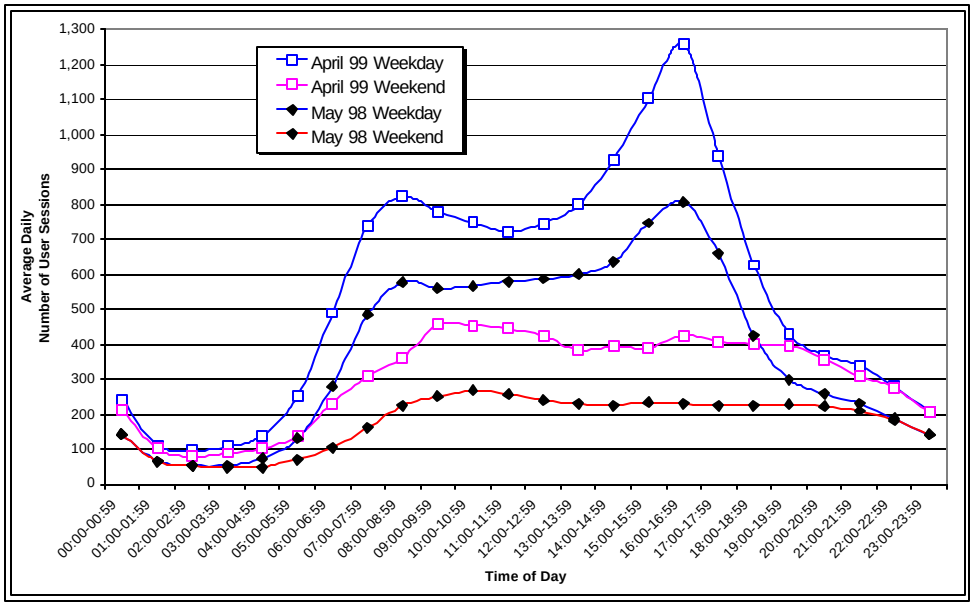


FIGURE 8. AVERAGE DAILY NUMBER OF USER SESSIONS FOR WEEKDAYS AND WEEKENDS, MAY 1998 AND APRIL 1999

traffic information from all sources during these two time periods. In addition, we know from focus group and survey research that commuters often experience more uncertainty and greater traffic congestion during the afternoon commute compared with the morning commute. We also know that more people have access to the Internet, and are more likely to use a computer for traffic information, at their office in the afternoon before leaving work than at home prior to departing for the office in the morning. Respondents whom we recruited from among WSDOT web site users generally reported more use of the Internet for traffic information in the afternoon. This is because of a perception that traffic is heavier then, and that there are more incidents in the afternoon than in the morning.[15] The average daily weekday use levels by hour for May and

[15] This perception is backed up by measured congestion on Seattle's freeways that shows a longer duration of congestion in the afternoon peak period than in the morning, a greater probability of encountering congestion in the afternoon, and slower average traffic speeds as a result (Ishimaru and Hallenbeck, 1999. *Central Puget Sound*

April (Figures 8) show a small peaking effect in the morning between 7:00 AM and 9:00 AM and a much more pronounced afternoon peaking in use in between about 2:00 PM and 5:00 PM. This typical bi-modal distribution, with separate peaks in the morning and a larger peak in the afternoon, increased dramatically over this year of observation. While we have not analyzed every month in Seattle at this level of detail, we have sampled several months to confirm that these general trends and patterns hold up for other months.

The weekend patterns of use show very little consistent variation by time of day, much as we would expect, given lower levels of congestion and no commuting. Comparing web usage on weekdays and weekends suggests that, while use is much higher on the weekdays, reflecting greater congestion associated with commuting, use levels on the weekend days is increasing proportionately about as much as on weekdays. Congestion in Puget Sound is getting worse over time and spreading throughout the day as well as throughout the entire week. A close examination of Figure 8 shows that web use in the mornings is starting out earlier and ending later in the afternoons in the April 1999 data compared with the May 1998 data.

Most Popular Page Views. During a user session, a user could view from one to many individual page views. Which pages they tend to view most reflects traveler preferences for certain kinds of traffic information, as well as effective ways of presenting and promoting that information. The web site pages with content that have value for the user will presumably be viewed more

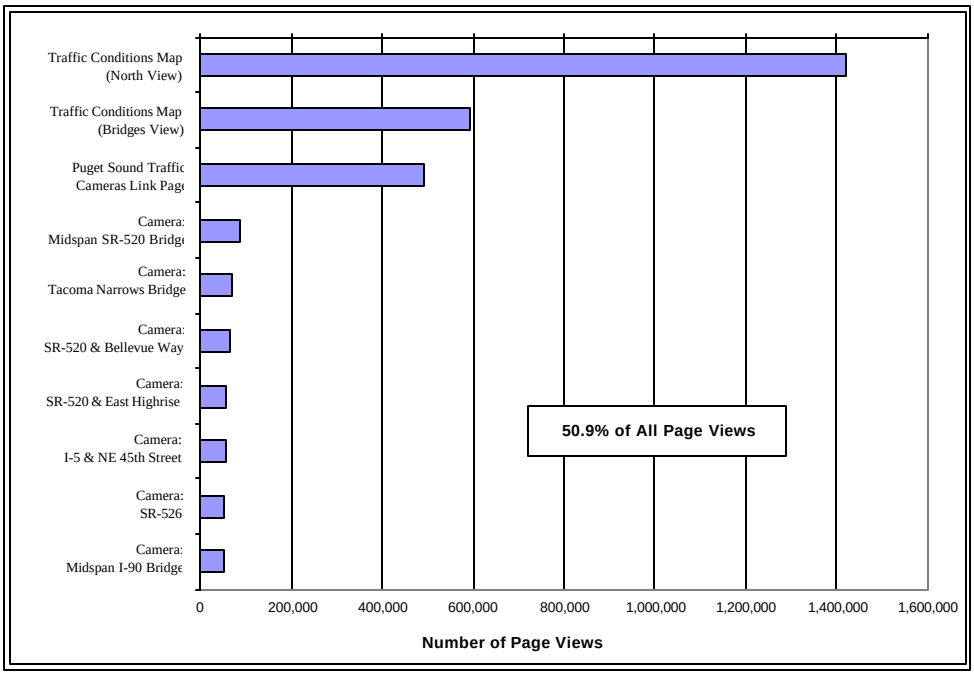


FIGURE 9. TOP TEN PAGE VIEWS ON THE WSDOT TRAFFIC INFORMATION WEB SITE: APRIL 1999

Freeway Network Usage and Performance. Draft Research Report, Research Project T9903, Task 62. Washington State Transportation Center, Seattle, WA.)

frequently. An analysis of the WSDOT web log files reveals the top 10 page views for the month of April 1999. This is illustrated in Figure 9.

Figure 9 shows that traveler interest in the WSDOT web site is concentrated in the top three traveler information pages, followed by a series of individual camera views. These three pages alone accounted for over 43 percent of all pages viewed in April. The top ten page views accounted for 50.9 percent of all pages viewed. The Seattle current traffic conditions map (North view) was the most viewed page (22.4 percent of all views), and the bridges view of this map was the second most viewed image (10.2 percent of all views) (see Figure 2 as well for an image of this map). Eight of the top ten page views involved camera images, either specific camera views or the main page where different views may be selected. The WSDOT site offers about 200 individual cameras showing traffic conditions on different segments of the freeway system in Seattle, and these are very popular with travelers. The Washington State ferries information page came in eleventh in rank order of pages viewed, with 0.9 percent of all page views for the month of April.

In addition to looking for the most heavily accessed web site page views, we can examine the relationship between page views and user sessions. The ratio of page views to user sessions by day for the months of May 1998, December 1998, and April 1999 is shown in Figure 10. The average for May 1998 is 13.7 page view accesses per user session. In December 1998 the average ratio was 15.2 and in April 1999 it was 16.5. This trend shows a steady increase over this period indicating that on average persons visiting the WSDOT web site are looking at two to three more pages each user session than they were a year earlier.

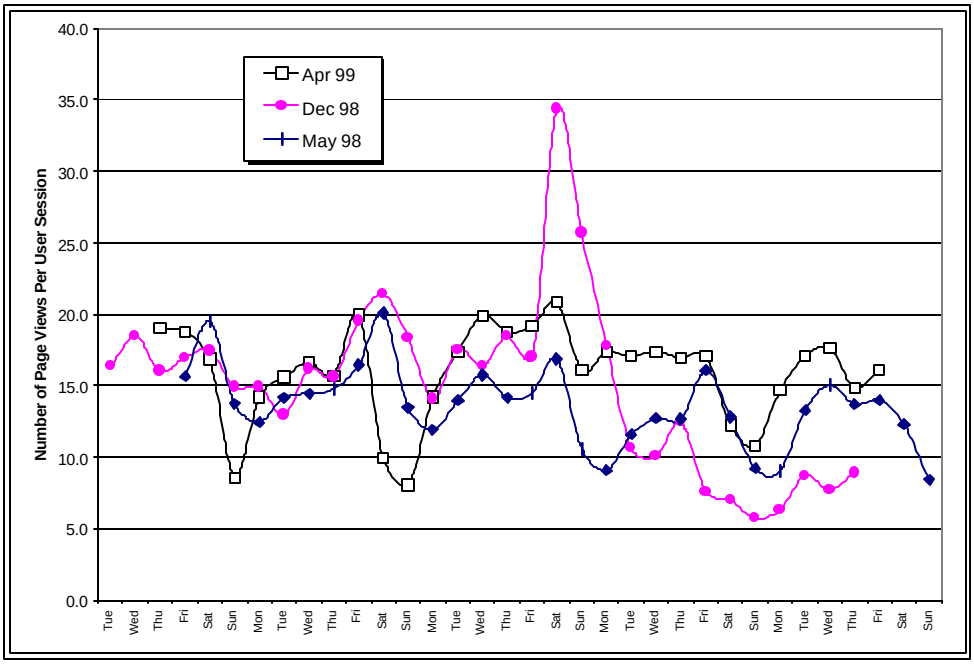


FIGURE 10. NUMBER OF PAGE VIEWS PER USER SESSION MAY 1998, DECEMBER 1998, AND APRIL 1999

Duration of User Sessions. We know from prior studies that many travelers have a high value of time; that is, they are in a hurry and don't want to spend a lot of extra, non-productive time during their busy days. While the acquisition of traveler information is valuable to them, we can presume that they want to get that information quickly, decide on their travel plans, and then get on with their day. WebTrends™ allows us to examine the length of user sessions. Figure 11 shows the length of user sessions in minutes for May 1998, December 1998, and April 1999.

The measure of user session duration by the WebTrends™ software involves a complex calculation that serves better for comparative purposes, both over time within a site, and between sites, than it does as a reflection of the actual amount of time that most users are likely to spend to acquire traveler information from these web sites. See Appendix B for a more detailed discussion on this issue.

The average user session in May lasted a little less than 24 minutes (23.9 minutes), and the average session in April lasted a little less than 28 minutes (27.9 minutes). We might expect that

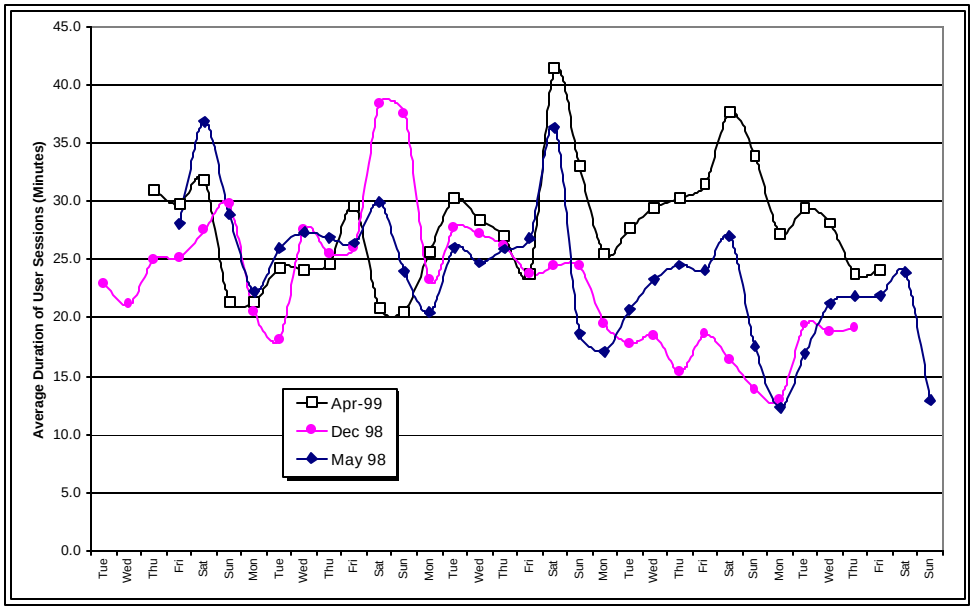


FIGURE 11. AVERAGE DURATION OF USER SESSIONS IN MINUTES, BY DAY MAY 1998, DECEMBER 1998, AND APRIL 1999

for sessions involving more page views that the session would last longer, and in fact the duration of user sessions is positively but only weakly correlated with the number of page views per user session in both May 1998 ($r = 0.18$) and April 1999 ($r = 0.07$). What is most striking about Figure 11 is the similarity in the patterns of user session duration throughout the period of observation. While the number of user sessions has increased substantially, the average number of pages viewed in each user session (Figure 10) and the amount of time spent by the user in each visit to the WSDOT web site (Figure 11) has remained relatively constant. In addition, there is some evidence of patterning by day of the week in Figure 11, with durations tending to be shorter during weekends and longer during weekdays.

Several factors can confound these data. For example, the transportation and safety agencies in the region are users of these data, and their use patterns may be very different from the average traveler, especially during periods of unusual weather, major construction, or other severe traffic disruption. Also, some users may keep their computers connected to the site for long periods of inactivity, only occasionally interacting with the site. Automatic processes such as server "pushes" of refresh information can keep the user session alive.[16] For these reasons, it is difficult to attribute some of these use patterns directly to end user preferences and behaviors. At best the web log data are suggestive, and can usefully be examined along with other data sources, such as focus groups and user surveys.

[16] See Appendix B for more discussion of these effects.

CONCLUSIONS

We can draw a number of tentative conclusions from this web use analysis:

- The level of use has grown significantly over the period of observation, doubling on average over this one year period of observation, and can be expected to continue to increase in the future. It is reasonable to anticipate that, as the miles of roadway under surveillance increase over time with further expansions of the WSDOT program, as new capabilities are added to the web site, and if congestion continues to intensify, usage of the site will increase. This will be reinforced by increases in computer use, Internet access, and awareness of the availability of WSDOT's traveler information web site.
- Though the underlying trend is clearly up, the WSDOT web site is characterized by month to month fluctuations of use of the site for traffic information. Two of the months, December 1998 and March 1999, showed significant jumps in usage compared to the average of their adjacent months of about 45% and 37% respectively. We have shown (Figures 3, 5 and 6) how unusual weather likely accounts for some of this increase in December, but we have not tried to analyze these "spikes" in further detail. The important lesson from these patterns of use is that latent demand for traffic information on the web can be significant; hence, to meet the needs of all users, it is important to plan to accommodate more users than visit the site on a normal daily basis. Information emerging from the analysis of the on-line WSDOT web survey suggests that users experience a slowdown in server speed and performance during periods of heavy use. In order to retain users and attract new users, server performance is a critical factor.
- Commuters are a significant component of any urban web site user population. The effect of their use is usually reflected in a pronounced peak in use during the afternoon commute period on weekdays and to a lesser degree during the morning commute period. This pattern substantiates other data that indicate greater Internet access at the office than at home, and heavier congestion with more uncertainty over a broader period of time in the afternoons. The WSDOT use data strongly support this pattern, reflecting the effects of severe congestion on commuter travel in Puget Sound.
- As is the case with other traffic web sites, travelers have a strong preference for dynamic traveler information over other static forms of information, including frequent updates on traffic flow and speeds, along with real-time camera views of traffic conditions. Eight of the top ten page views on the WSDOT site recorded in April 1999 involved camera or video-related material. Usage is spread out over more pages on this site compared with most other traffic information web sites. WSDOT includes more dynamic information pages (200) compared with the other MMDI sites, and use patterns indicate that Seattle area travelers are accessing and value this broad array of available information.
- Recently, WSDOT users have been visiting an average of about 17 different pages during a user session. This represents a small increase in the average number of page views per user session from about 14 per user session a year earlier, perhaps reflecting the addition of more useful pages, including new camera images. Further research will be needed to better understand this apparent trend.
- Our examination of the duration of user sessions over this 12 month period also shows an increase in the mean duration of user sessions from about 24 minutes to 28 minutes. This is a longer duration compared with observations at the other MMDI sites. However, it is difficult

to interpret the duration figures, since a number of external factors and assumptions can impact the length of a user session (see also Appendix B).

Traveler information web sites are becoming very popular with state DOT's around the country and Seattle's web site is no exception. The infusion of MMDI funding in support of better and more information that can be provided through the Internet is being well received by travelers in the Seattle area. This is reflected in the rapid growth in the use of the site over the past year. WSDOT is actively developing their WSDOT web site and expanding the range of dynamic traveler information offered through the Internet. The evidence suggests that this will continue to be of significant benefit and well received.

Appendix A

Management of Web Use Log Files

Web servers generate log files that track different kinds of visitor activity at a web site. The activity types contained in the log include the pages visited, how long the visitor or user session lasted, the user path through the site, and other information. These data, preserved in the log file, can be analyzed to help determine how users actually make use of the information presented on the web site.

Log File Acquisition

Log files can be set by the webmaster to measure activity for a day, a month, other preset time periods, or a specific range of time. Because of the busy nature of some web sites or the length of time set for the log file to remain active, the file size can become quite large. WSDOT compiles their web use log files into separate one month time periods.

Battelle worked with the WSDOT webmaster, Mr. Wayne Szydtowski, and Mr. David Richards, to acquire copies of these log files. File copies were placed on an ftp server accessible from outside the WSDOT system. Battelle accessed this server and copied the files to a server in Seattle. These files were then permanently recorded on a CD for storage and subsequent off-line processing and analysis.

Log File Analysis

Several commercial software products have been developed for web server log file analysis. The product used by Battelle is manufactured by WebTrends Corporation (http://www.webtrends.com/default.htm). WebTrends™ Professional Suite contains components for analyzing web site activity, link analysis, and other tools for managing web sites.

WebTrends™ Professional Suite comes with default analysis formats that can be modified by the customer. The default settings contain mechanisms for including general statistics (timeframe, total hits/page views/user sessions for site, average number hits/page views/user sessions per day, etc.), most requested pages, top entry/exit pages, most downloaded files, most active organizations, most active countries, activity level by day of week/hour of day, client/server errors, top referring sites, most used browsers, and most used platforms.

For the MMDI analyses, WebTrends™ was customized to process the elements identified as probably most useful for understanding use patterns at transportation web sites. These elements included general statistics, most requested pages, top entry/exit pages, single access pages, top paths through the site, most active organizations, most active countries, summary of activity by day, activity levels by day of week/hour of day, client/server errors, North American States and Provinces, active cities, top referring sites, top referring Universal Resource Locators (URLs),

and top search engine information. Not all of these elements turned out to be data types actually used in the overall analysis.

There were many interesting problems that emerged during this acquisition and analysis process. One example is that some monthly log files lapsed over a few seconds or minutes into the next month. WebTrends™ averages were based on the number of days contained in a particular time period. If 'month' was selected in the setup rather than specifying a date range, WebTrends™ computed averages that factored in an extra day, even though that day registered for only a few seconds. For those analyses, the averages were recalculated to get the correct figure.

Another problem experienced was in the Most Requested Pages category. Log files would sometimes register a single page as two or more different pages depending on how the server viewed the URL request from the user. In these cases, the WebTrends™ analysis had to be revised, and the multiple page totals were combined when the page was in or close to the Top 10.

A similar problem occurred when analyzing sites with frames pages. The log registered the main frames page and the component subpages as individual pages. Because these are all really a single page, the analysis had to be reviewed manually and recompiled.

Initially we believed that monthly reports would be sufficient for web use analysis. Exploratory data analysis at several sites suggested that it would be necessary to do a separate analysis for each day in those months that exhibited wide swings in use patterns in order to understand what might be causing the variance. In addition, daily WebTrends™ analysis proved to be the only way to examine daily and hourly patterns at a level of detail below monthly averages. This was then followed by seeking external information to help explain the causes of observed patterns in daily use variation.

Appendix B

Measuring the Number and Duration of User Sessions Using WebTrends Ô

WebTrends™ provides the ability for the user to designate the length of a time interval that will be used to define a user session. When a user IP address is first detected in the web log file, a new user session is begun. Subsequent instances of that same IP address in the log file sequence are included in the user session as long as the designated time interval between one instance of that address and the next is less than the specified interval. Once that interval is exceeded, due to a lack of activity, the user session is considered terminated and is set equal to the length of time between the first detection of the IP address and the last just prior to the time interval that exceeds the specified default value. Alternatively, if the user exits the site before the interval is exceeded, the user session is terminated at that point.

There are two ways a user's IP address can appear in the log file. One way is for the user to actively seek out information, connect to the site, and visit one or more pages or picture images. The other way is for the server to automatically push information out to the user's computer after an initial connection to a web page. This takes the form of automatically updated or refreshed information related to the dynamic information pages on the site, such as the flow map or the camera images. Thus, the server will create information associated with an IP address in the log file, even without any active intervention by the user, other than the user's decision to stay connected to that site. Since these server actions can occur as frequently as every minute, a default interval longer than that would never be achieved and the user session would continue until the visitor logged off the site.

The analyses conducted for the WSDOT traffic information web site uses the WebTrends™ default value of 30 minutes to define user sessions. We can hypothesize that there are several patterns that describe user sessions for this kind of traffic information. Many visitors to this web site can be expected to spend a very short time, perhaps no more than a couple of minutes, acquiring the specific information they need to make a travel decision. After they get that information, they would log off. Their user session would be short. Others may visit the site to see what traffic conditions are like and remain connected to the site for a period of time, checking traffic status from time to time. When conditions seem acceptable, then they would decide to log off and leave. Their user session length could be quite a bit longer. Finally, we know from focus group discussions that there are some users who leave their computer connected to this site for hours at a time, allowing for casual inspection of traffic conditions from time to time. This results in a few potentially very long users sessions that can skew the overall average.

The length of the WebTrends™ user session interval has a very significant effect on the calculation of the number of user sessions, as well as on the duration of a user session. The longer the interval of inactivity specified, the fewer the number of user sessions that will be calculated and the longer the duration for those user sessions. Conversely, the shorter this time interval, the greater the number of user sessions and the shorter their duration. We conducted a sensitivity analysis by systematically varying the length of the time interval and observing how

many users sessions were identified and the duration of those sessions in minutes. Figure B-1 illustrates the results.[17] For intervals set at 30 minutes (the default value) or longer, the number of sessions identified declined gradually and the duration of those sessions lengthened gradually. For intervals set less than 30 minutes, the number of user sessions identified increased very dramatically and correspondingly the duration of those sessions dropped dramatically.

There is no "right" value to specify for this time interval in WebTrends™. We have chosen to stay with the default value of 30 minutes for this analysis for several reasons.

- This is the value that is used by the state DOTs in defining user sessions in the analyses they typically make available to the public. To change this value would create unnecessary confusion and make it more difficult to interpret the findings on web usage across different analyses of the same data or across sites.
- As shown in Figure B-1, the number of user sessions identified tends to stabilize with a value of 30 minutes and change relatively little with longer intervals, while the results change substantially for shorter intervals.
- As the interval value is shortened below 30 minutes, the risk increases that legitimate single user sessions will be arbitrarily truncated. Thus, for very short intervals, a single session with one person examining traffic information periodically might be represented as many apparent individual sessions due to setting too short an interval of inactivity at that site.
- As we have indicated earlier in this report, user sessions are at best a proxy for individual users, and WebTrends™ doesn't allow us to clearly identify every user, or to distinguish

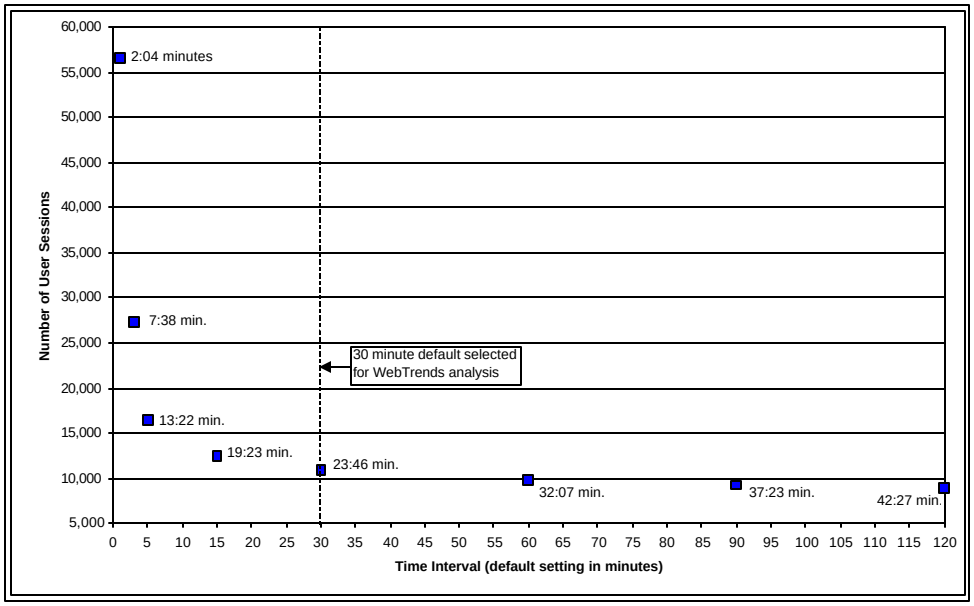


FIGURE B-1. SENSITIVITY TEST: DEFAULT TIME INTERVAL BY USER SESSION LENGTH AND USER SESSION DURATION (WSDOT WEB LOG FOR 4-29-99)

[17] This analysis was conducted on the WSDOT daily log file for April 29, 1999, a normal-appearing weekday in the most recent month covered in our analysis. We have not conducted a more extensive sensitivity analysis to

repeat users among multiple user sessions. The number of user sessions derived from the log data by WebTrends™ is the best indicator we have to reflect user activity levels at the web site. By keeping the same interval default value throughout the analysis of one site and at each of the other sites, relative changes in the levels of usage and changes in the patterning of use can be clearly identified.

determine whether there is variation in these results for other days or other sites, though this would be a useful next step.