ITC 16 / P. Key and D. Smith (Editors)

A preliminary analysis of Work-At-Home Traffic Characteristics

P. Orenstein

Bellcore, 331 Newman Springs Road, Red Bank, NJ 07701, USA *

ABSTRACT

According to the 1997 Residential Telecommunications Survey, 25 % of US households are working at home in some fashion. With the continued interest in flexible work schedules and work-at-home possibilities, this population segment is likely to grow and exacerbate congestion problems for the Public Switched Telephone Network (PSTN). For the foreseeable future, the PSTN will provide the majority of users with access to the Internet, in addition to 'regular' telephone or fax x traffic. This represents an enormous shift in the volume and nature of the PSTN traffic.

Until now, traffic data studies have concentrated on the distinguishing features of online[1] connections from voice calls. It is now accepted that the average holding times of online connections can be many times the holding times typically assumed in voice telephony and that the peak usage of online traffic does not coincide with the conventional mid-morning busy hour associated with voice traffic. However, data studies about online traffic have focussed mainly on Internet traffic[2], and information about the nature and characteristics of other forms of online traffic is sparse.

This paper seeks to explore the general characteristics of work-at-home traffic, and compare them with corresponding features of voice traffic and Internet traffic. The work is preliminary in nature and is based on a limited sample of users. However, it highlights the main causes for concern for network engineers and planners by providing an in-depth analysis of call hold times, busy hour periods, and subscriber behavior which to date has not yet been undertaken.

---

* email: porenste@notes.cc.bellcore.com

[1] For the purposes of this discussion online traffic may be taken to include ISP, work-at-home (WAH), telecommuting and online service calls, all of which appear to have common characteristics.

[2] By Internet traffic we mean subscribers to Internet Service Providers (ISPs) specifically for access to the Internet

1. INTRODUCTION

The past year has seen explosive growth [1] in online related telephone traffic – specifically, calls from residential and business customers across the PSTN to Online Service providers. Whilst the increase in volume poses an immediate threat to the capacity of the PSTN, at a more fundamental level, its qualitatively new characteristics are challenging the engineering, forecasting and planning procedures established by the former Bell System over the past 80 years.

Today's PSTN is an efficient carrier of voice traffic. It has evolved through an emerging understanding of the characteristics of voice traffic, which adhere to the following well-established assumptions:

a) The average call holding time is around 3 minutes[3]
b) The statistical call holding time distribution is well approximated by the exponential distribution, and
c) Call arrivals in the "busy" time periods in which "memoryless" behavior is obtained follows a Poisson probability distribution

These mathematical assumptions have been validated via analyses of measured data. These assumptions formed the basis for a detailed yet fundamental study [2], which compared and contrasted Internet traffic and voice traffic. The study highlighted the importance of studying the characteristics of Internet call data and also provided initial PSTN trunk and switch engineering algorithms for coping with Internet long holding time traffic. This led to more extensive work into developing architectural solutions to PSTN congestion based on intelligent-network-capabilities [3].

But, the need to tailor Internet offload strategies and modify existing engineering algorithms to keep pace with the changing characteristics of network traffic, has become critical with the rise of online traffic. A recent market survey supports this claim [4]. It suggests that a growing percentage of households are working at home in some capacity. In 1998, 27.5 % of U.S. households made up the WAH category – a 10 % increase from last year. Moreover, a need has arisen to develop a more comprehensive description of PSTN traffic which is not confined to "busy-hour" periods but which extends to an entire 24 hour period and sometimes beyond. The purpose of this paper is to describe the behavior of "online traffic" throughout the day with a particular emphasis on WAH traffic. In addition, we highlight the differences between conventional Internet traffic and WAH traffic and discuss the implications for network planners and coordinators arising from these differences.

2. BACKGROUND TO THE DATA STUDY

The community of Work-at-Home users is usually partitioned into three groups. These include (a) telecommuters, (b) corporate after-hours and (c) self-employed. The

[3]This figure is an historical consumer reaction to specific AT&T (Bell System) tariffs

data source for this study comes from classes (a) and (b). The data has been gathered from Bellcore's own corporate Intranet over a three-month period in 1998. The collected data has been processed into call records. Each record corresponded to a single call made by an employee who was accessing his/her computer account remotely. Employees dialed up to their accounts remotely either by using a dedicated 'toll-free' telephone-number, or by selecting a local access number. More recently (Sept 98), a digital service was also offered to employees although the penetration of this service was limited to ISDN users only. The data set consisted of a total of 92008 'toll-free' calls, 93115 'local-access' calls and 4938 calls from the 'digital' service. The data was divided into weekday and weekend data.

The statistical analysis is preliminary in nature and relates to a relatively small community of users (Bellcore's workforce is currently about 6500 employees). Of these about 50% used the remote-access service in some form. The analysis points out the differences between the various forms of online traffic using the PSTN.

The first part of the study investigates general statistical characteristics of the data (eg average call lengths, cumulative probability distributions, peak busy-hour periods, daily usage patterns and individual subscriber behavior) for both 800[d] and non800 users, focussing particularly on weekday traffic. The second part of the analysis takes a sub-section of the weekday data only and compares the WAH daily usage patterns with different types of traffic data (eg ISP traffic), in order to assess the combined impact and changing characteristics of different types of network traffic accessing the PSTN.

3. A CHARACTERIZATION OF WORK-AT-HOME TRAFFIC

3.1 General characteristics

The first part of the analysis partitioned the data set into three groups: toll-free, local access and digital. Each of the groups was then subsequently divided into weekday and weekend data. Table 1 summarizes the average call lengths measured in minutes.

From the statistics shown, one can observe that the average call holding time for WAH users is in the range of one hour, although typically, the average call exceeds one hour. The digital service had the lowest average weekday CHT of 47.71 minutes. By contrast the average CHT for 800 users was almost one and a half times longer and for non800 users almost two and half times longer. This could be due to the fact that those using the digital service were gaining faster access because they all had ISDN capabilities. This may have reduced the actual time spent online.

Another feature of these statistics is that the mean CHT for weekend traffic is lower than the corresponding weekday traffic. This is also true of the "engineered" call volumes.

[d] In the U.S. 800 numbers are 'toll-free'

Table 1
CHT Statistics for WAH traffic

	Total Call Vol. K(min)	# Calls	Mean	Std Dev
800 wday	5194.25	76431	67.96	174.26
800 wend	903.93	15577	58.03	151.45
Non800 wday	9367.27	78545	119.26	425.81
Non800 wend	1509.59	14570	103.61	445.02
Digital wday	52.24	1095	47.71	125.43
Digital wend	25.87	450	57.51	90.01

3.2 Distribution of Call Holding Times

Having looked at average call hold times of WAH traffic, we now consider the distribution of these calls. We have only taken the weekday traffic since this was the more substantive data set. The cumulative probability distribution curves for the different services were generated. The results are shown in Figure 1.

The distribution of the three services across the group of remote-access users can be summarized as follows. Approximately 26 % of the group were dedicated 800 users and similarly 26 % were found to be dedicated 'local-access' users. There was a sizable overlap between these two groups. Roughly 45 % of users used both the 'toll-free' service or the 'local-access' number. The remaining 3 % used the 'digital' service.

The comparative cumulative distribution curves for the different services highlight the differing characteristics of network traffic. The '800' customers displayed consistently lower holding times than their 'non800' counterparts. One explanation for this is related to the charging pattern in use at Bellcore. Whilst the 'local-access' number incurred no cost to the user's organization at Bellcore, (providing the users were truly local), the toll-free number did incur a time-based charge to the organization. Consequently, when a user called the 'local-access' number, they would remain connected for longer periods, since the call was absolutely free, which was not the case for the 800 service. The digital service shows the shortest CHTs of the three services. This service is the fastest of the three and is the most reliable. However, its penetration is low compared to the 800 or non-800 service.

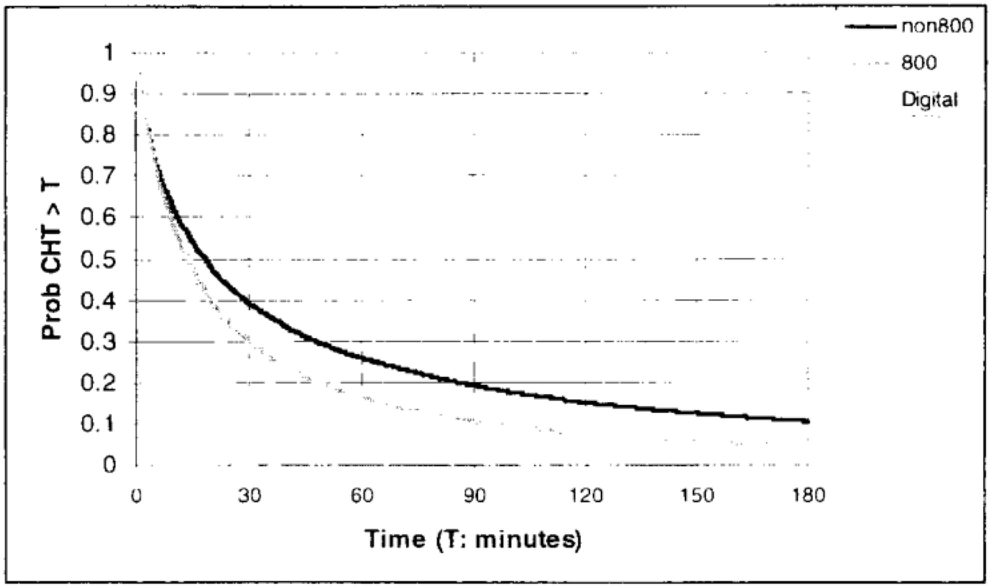

Figure 1. Comparative Cumulative Probability distributions for WAH traffic.

3.3 Daily Usage Profiles

Next we describe the daily behavior of WAH traffic. By partitioning the weekday data from one month (September) into separate days and observing the distribution of usage over a twenty-four hour period for each day we were able to highlight the proportion of telecommuters and corporate after-hour traffic accessing the corporate Intranet and assess their behavior.

The proportion of usage generated in each hour was measured for each of the different services (800, non800 and digital) on a daily basis. The average behavior over a twelve-day period was then calculated and is shown in Figure 2.

Telecommuters vs. After-Hour Corporate users

From the graph one can observe the two types of WAH traffic: telecommuters who access the network during the standard workday business timeframe of 9AM-17PM, and corporate after-hours traffic who continue to use the service outside the regular 'business' hours. The graph shows how the proportion of usage generated by telecommuting traffic grows during the standard workday hours, and reaches its peak of approximately 8 % of the total daily usage between 11PM-12PM. This will be referred to as the *telecommuting busy-hour*. Following a decrease in the load, WAH traffic begins to pick up slowly during the mid to late evening hours. A second, less noticeable peak occurs at 22PM-23PM. This corresponds to the *corporate after-hours busy hour*. From the data, it is apparent that most of the 800 and non800 users form the bulk of the telecommuting traffic and do not contribute as much to the corporate after-hours traffic. In sharp contrast are the users of the 'digital' service who appear to contribute less to the telecommuting traffic and more to the 'after-hours' corporate traffic.

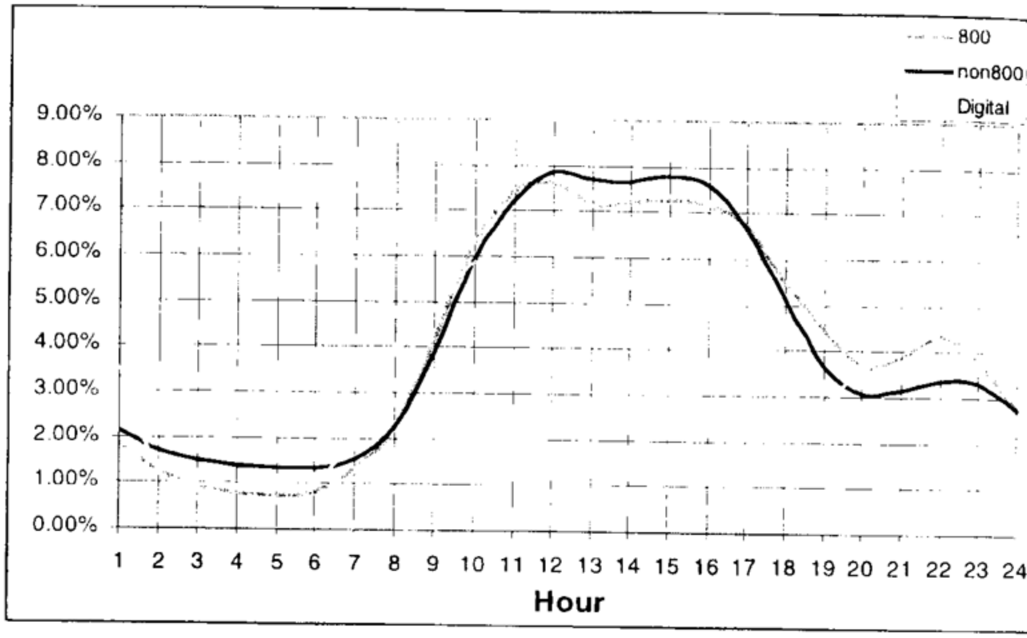

Figure 2. Average daily percentage of usage for Sept 98

4. WAH TRAFFIC – A COMPARISON WITH OTHER FORMS OF NETWORK TRAFFIC

Until now, WAH traffic has usually been associated with ISP traffic since it was accepted that both types of traffic shared common characteristics. However, no comparison has been made so far using actual data. We next compare WAH traffic with conventional voice traffic and ISP traffic and point to the common features as well as those that distinguish between them.

To make the comparison, we used data from a central office in Northern New Jersey, which was in the geographic location of the WAH data. We extracted voice calls and calls made to a dominant ISP in this region from the central office data for the month of September 1998, (the same month for which this part of the WAH analysis was based).

Table 2 and Figure 3 show how network traffic can span a very wide variety of time scales: the average CHT for voice traffic is 4.5 minutes, for ISP traffic 21.35 minutes whilst for WAH traffic (with no differentiation between '800' and 'non800' calls), the average is 94 minutes. In addition, the standard deviation for voice traffic is relatively low indicating that voice calls are fairly uniform in nature. In contrast, there is much variability in both ISP calls and WAH calls as shown by the large standard deviation figure. Although both ISP and WAH traffic exhibit long holding time patterns, Figure 3 shows that the overall distributions are not the same. For instance, the probability of voice call exceeding one hour's duration is less than 1%. In contrast, more than 5% of Internet calls will exceed one hour and over 20% of WAH calls last for over one hour. Furthermore, the tail of the WAH distribution is much longer than that of ISP traffic. This is likely to present some "engineering" difficulties for network planners.

Table 2
Average Call Lengths for different types of network traffic

# Calls	Mean CHT (Min)		Std Dev
Voice	35000	4.48	2.68
ISP	268066	21.35	52.85
WAH	154976	93.96	326.6

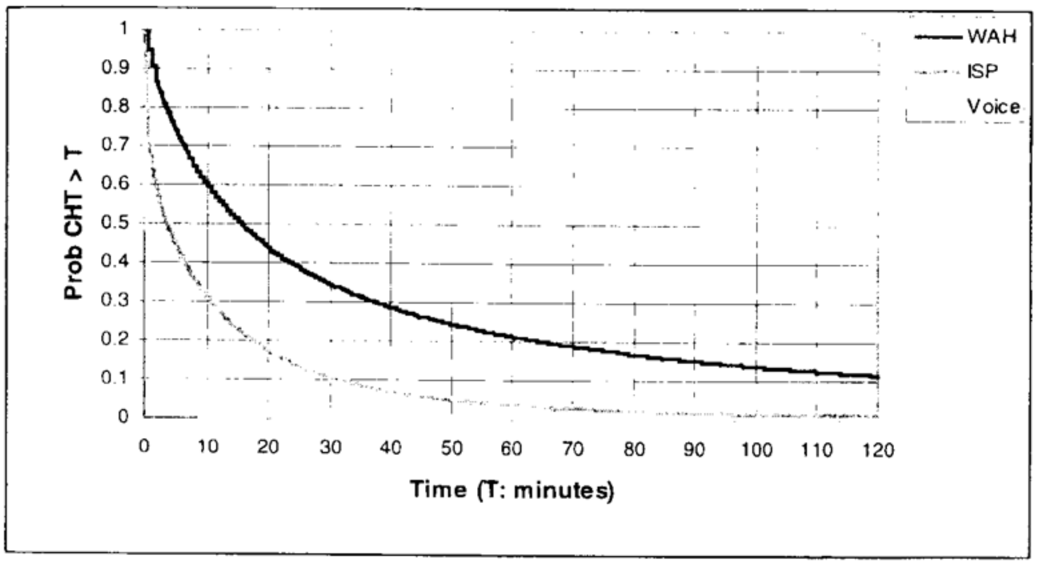

Figure 3. A comparison of CHT distributions for various types of network traffic

To understand why the explosive growth in WAH traffic is likely to cause some engineering difficulties, we compared the proportion of daily usage for voice, ISP and WAH traffic as distributed over a single day in September. Figure 4 shows the superposition of the distributions of traffic usage over a twenty-four hour period for the three types of network traffic. (Similar distributions were obtained for other days in September).

From the graph, one can observe how the pattern of usage for WAH traffic seems to be reflected in voice traffic. The *telecommuting busy-hour* roughly coincides with the voice busy hour (11AM-12PM) and a second voice busy-hour can be seen at the same time as the *after-hours corporate busy hour* (21PM-22PM). ISP traffic is a different story. The usage follows a steady growth pattern throughout the day with the ISP busy hour occurring at 22PM-23PM. Moreover, the % of usage generated in this busy hour (11.5%) is much higher than that of the voice or WAH busy hour periods.

The superposition of the three graphs explains why the growth in network traffic will have a critical impact on the PSTN's resources. Before the rise in 'online' traffic, long holding time traffic did not interfere as much with voice traffic since their respective busy hours occurred at entirely different times in the day. But, with the rise in popularity of WAH traffic and the continued expansion of ISP traffic, long holding time traffic is likely to impact the PSTN throughout the day. This means that the boundary between the ISP busy hour and the conventional voice busy hour will become blurred and the PSTN will need to deal with excessive call volumes and long holding time traffic for more extensive periods of the day.

5. USAGE CHARACTERISTICS OF WAH TRAFFIC BY SUBSCRIBER

Having explored the general features of WAH traffic we now evaluate the individual behavior of subscribers using the dial-up service. We have used the weekday data from the month of September for both 800 and non-800 users (a total of 17 consecutive weekdays or just under 4 weeks of weekday data). In the data sets, each call record was associated with a single user.

An important issue which arises in LHT traffic is what proportion of the user group is generating the majority of the traffic load. In other words, is there a small group of 'persistent' users who are responsible for a large portion of the load, or is the load evenly distributed among all users? We refer to the diagram shown in Figure 5, in which the cumulative distribution of WAH usage Vs WAH users is shown.

The graph highlights the difference between users of the 'toll-free' (800) service and users of the 'local-access' (non-800) service.. Over a period of 17 weekdays approximately 7 % of the 800 users generate 50 % of the total traffic usage vs. 2.9 % of non-800 users. Similarly, less than 1 % of the non-800 users are responsible for a quarter of the traffic load vs. just under 2 % among the 800 users. This shows that there is a core group of 'persistent' users in each group who generate a substantial portion of the WAH traffic, although the group appears to be smaller in the '800' case.

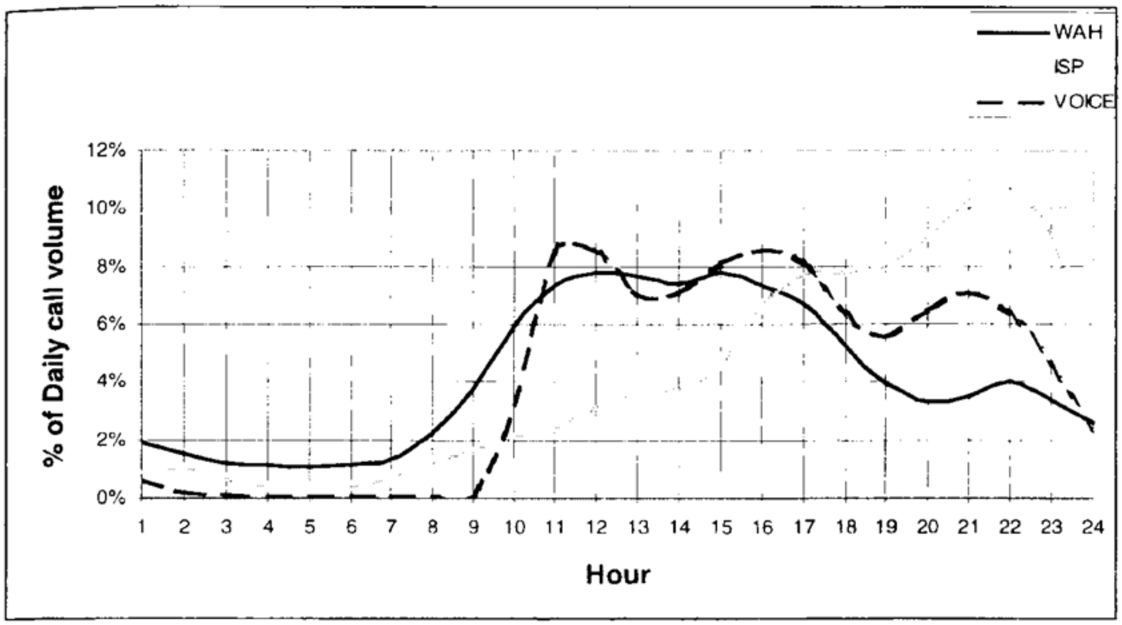

Figure 4. Distribution of call usage over 24 hours for different types of network traffic .

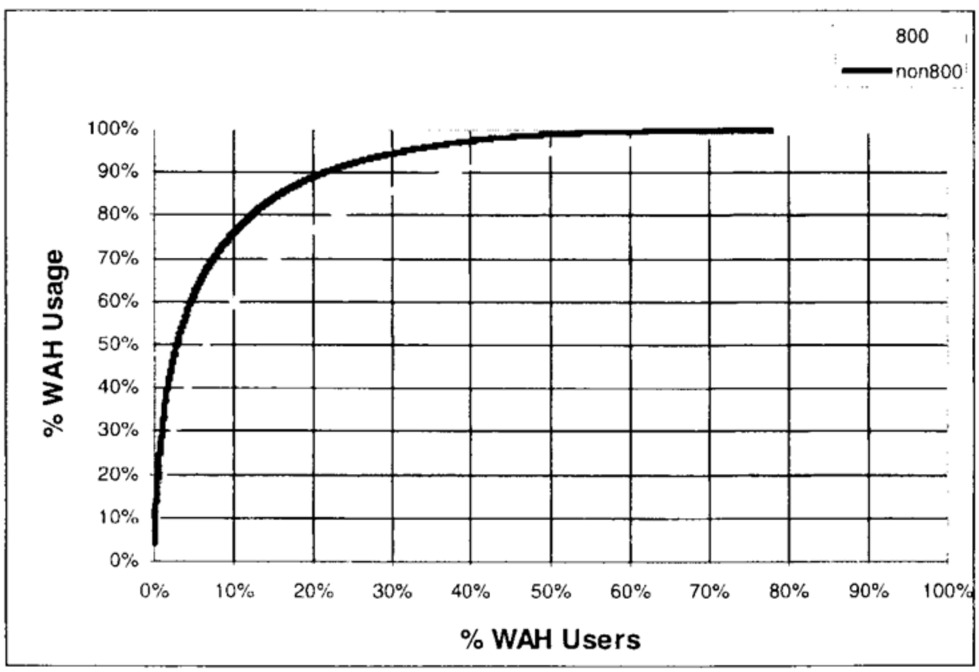

Figure 5. Heavy Usage Curves for WAH traffic

This leads us to consider the nature of the different types of users. Essentially, there appear to be three groups of users: persistent, frequent and occasional. Persistent users are those subscribers who use the service on an almost daily basis. Frequent users access their accounts on a regular basis, typically once or twice a week. Those who subscribe to the service on odd occasions have been termed occasional users. Figure 6 classifies subscribers by the total number of days they are active. The graph shows the cumulative distribution of the number of active days vs. the proportion of subscribers who are active. The curve shows the proportion of subscribers decreasing as the number of days increases. This trend occurs for both '800' and 'non-800' users. Persistent users were in the top 10% of the sample and recorded usage on at least 12 days. Frequent users recorded usage on between 2 and 11 days (approximately 60% of the sample). The remaining 30 % of the sample of WAH users recorded usage on one day out of 17. These were classed as occasional users.

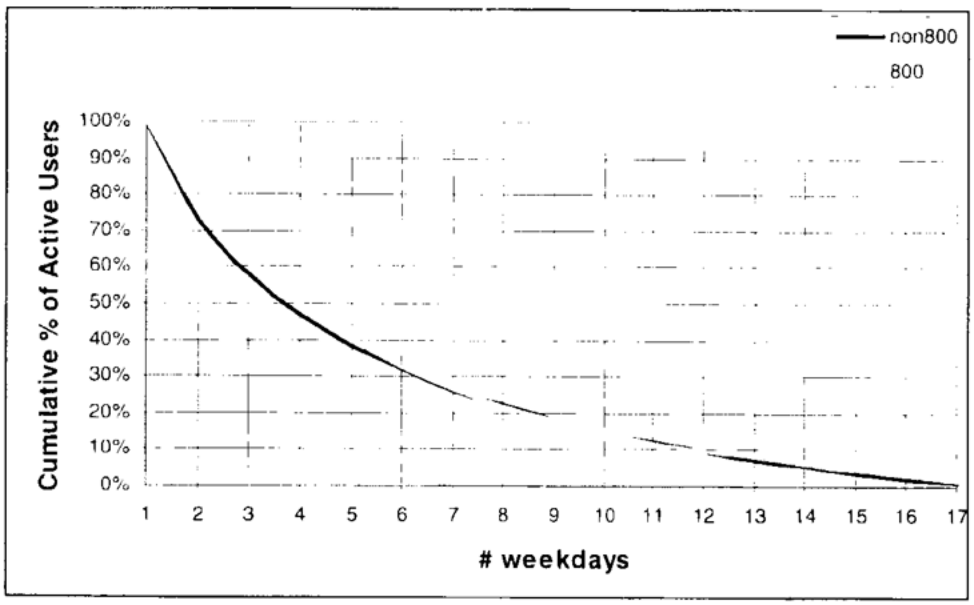

Figure 6. Proportion of active subscribers by number of days they were active

With respect to the proportion of usage generated by subscriber, we have observed that relative to the proportion of users, both 'persistent' and 'frequent' users generate more traffic than subscribers who use the service occasionally do. Figure 7, demonstrates this graphically for the '800' users. A similar affect was noticed with the 'non-800' users. The solid black bars show the proportion of users who fall in a particular category. For example, approximately 3 % of the '800' users are active on all 17 days, whereas roughly 25% dial in only once during this period. The graph shows that the proportion of 'persistent' users is restricted to a small group. By contrast, the usage they generate is relatively high: for example, the 3% who are active on all 17 days are responsible for just under 15 % of the total usage, whereas the occasional subscribers (25 % who are active on 1 day only) generate only 3 % of the total usage. This emphasizes the different types of behavior in the WAH community.

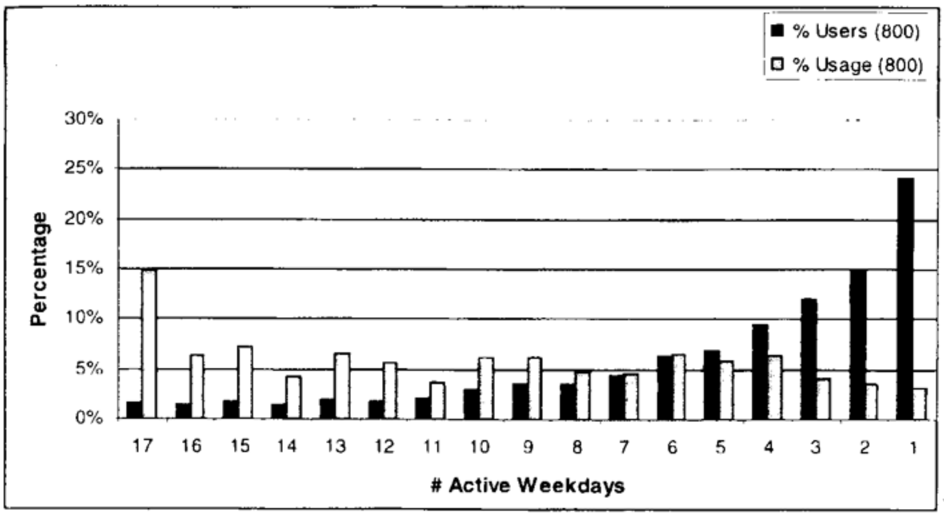

Figure 6. Proportion of Usage generated by different classes of WAH users

6. CONCLUSIONS

Since the PSTN currently represents the only near-universal access method for 'online' connections, it is important to continuously evaluate the impact of changing network characteristics through studies of measured data, in order to develop strategies to help ease congestion on the PSTN. However, whilst the Internet phenomenon is relatively well established, less is known about the Work-At-Home (WAH) scene. This paper provided an in-depth analysis of the general characteristics of WAH traffic and compared them with other forms of network traffic by examining three months of WAH data gathered from Bellcore's corporate Intranet. The main findings can be summarized as follows:

Two types of WAH users were identified in the data: telecommuters who were active during the 9AM to 17PM timeframe and after-hours corporate users whose main usage occurred in the mid- to-late evening hours. The busy hour for telecommuters was between 11AM and 12PM whereas for the after-hours corporate users, the busy hour was at 21PM-22PM. This roughly coincided with the ISP busy-hour, although the average CHT for ISP traffic was typically in the range of 21 minutes compared with over one hour for WAH traffic. For WAH traffic, users of the digital service exhibited the shortest CHTs whereas users of the 'local-access' service had the longest. Users of the toll-free service had an average CHT somewhere in between.

In the WAH sample there appeared to be three groups of users: persistent, frequent and occasional. Persistent users focussed on a small core group of users who were responsible for relatively high proportions of traffic. This was in sharp contrast to the large number of occasional users who contributed a much smaller proportion of the overall traffic load. The study emphasized the different types of behavior in the WAH community. A natural extension to this research would be to evaluate the differences between the forms of WAH traffic identified here, for example, subscribership issues for telecommuting vs. after-hours traffic and a detailed discussion of usage patterns among users in each group.

REFERENCES

1. H. Sander, Work-at-Home Households: A major telecom stronghold, IDC/LINK Bulletin No. 13555, Vol. 1, (1997).

2. J. Gordon, P. Murti and J. Wang, Engineering the PSTN for Internet Traffic, Bellcore SR-4109, (1996).

3. A. Atai and J. Gordon, Architectural Solutions to Internet Congestion based on SS7 and Intelligent Network Capabilities, Bellcore White Paper 00A-1019W, (1997).

4. M. Porter and R. Boggs, U.S. Home Office Market Review and Forecast 1997-2002, IDC/Link Bulletin No. 16648, Vol. 1, (1998).