Measurement and Modelling of Internet Traffic at Access Networks

Johannes Färber[1], Stefan Bodamer[1], Joachim Charzinski[2]

[1] *University of Stuttgart, Institute of Communication Networks and Computer Engineering, Pfaffenwaldring 47, D-70569 Stuttgart, email: {faerber|bodamer}@ind.uni-stuttgart.de*

[2] *Siemens AG, Public Communication Networks, Hofmannstraße 51, D-81359 München, email: Joachim.Charzinski@oen.siemens.de*

Abstract

With the growing popularity of the Internet, more and more users access the Internet from their homes via the public switched telephone network. Today´s traffic load can not be described accurately by classical methods and therefore new traffic models are needed to allow proper network design and dimensioning. In this paper the dialup session behaviour of home users is discussed. Based on the log data of the modem pool of the University of Stuttgart we present typical traffic characteristics and suggest a simple traffic description for use in traffic models on session level for performance analysis or simulation.

1 Introduction

The behaviour of Internet users has a major impact on the performance of networks and services. Especially telephone network operators are interested in descriptions of user generated Internet traffic to be able to adjust their current networks to the growing demand. Such traffic descriptions must be based on measurements of real traffic.

The modem pool at the University of Stuttgart offers the possibility to observe the users´ session behaviour in large scale and for a sufficiently long period. With around 4000 subscribed students and 800 subscribed staff members, the user population is large enough to draw conclusions on general user behaviour.

In this paper we present the results of an evaluation of the modem pool log data from May to October 1997 with 369100 sessions. The log data contain information on the login and holding times for access sessions. In contrast to many other publications concentrating on the packet level, we focus on this session level, i.e. the characteristics of the users´ dialup sessions without regarding the type and quantity of information transmitted during the sessions. In Section 2 we describe this observed traffic behaviour.

For performance evaluation of network components, traffic load has to be modelled under which the system performance can be measured. In Section 3 we present a simple modelling approach that allows to generate traffic with characteristics similar to the monitored traffic.

2 Session behaviour

The automatic monitoring of the user login times at the modem pool of the University of Stuttgart allows the evaluation of characteristic measures on session level. Since students and members of staff are assigned two separate modem pools, we distinguish between these two user groups. Note, that there is no way to specify the type of session the user has started. While in most cases it can be expected to be a World Wide Web session, it may also be a telnet session, an ftp or a simple email retrieval or a mix of those traffic types.

In the following sections we describe the holding time of the sessions, the interarrival time between session starts and the mean daily traffic profile for traffic load. These measures allow to characterize the frequency and duration of a typical user session.

2.1 Holding time

Under the holding time of a session we understand the duration of the seizure of a modem. The mean holding time was around 21 minutes for students and 20 minutes for staff members. However the holding time shows a high variability. When comparing the students of different subjects the observed mean holding time varies from 8 to 31 minutes. Single users have even been online for several days. The maximum holding time was 11 days.

The holding time varies strongly during the course of the day. Figure 1 shows the holding time of sessions (of both user groups) associated with the time of the session start. Although this representation has to be regarded with caution (the average during the night is calculated from a relatively small number of calls), it allows the conclusion that long sessions start mainly during the night and early morning hours. Also the mean session length at night is significantly larger than during the day time. The average sessions in the mornings or during the day of non-workdays are longer.

In [6] Morgan reports a significant peak in holding time at 4 am. If the holding time would have been drawn associated with the session endings we would have obtained a peak at 4 am as well. This means that among sessions that end in the early morning, most have lastet for a long time and only few are of a short duration.

The high variability of the holding time is visible in the complementary cumulative distribution function (CCDF) which is depicted in Figure 2. The function shows the probability of a holding time being greater than the value on the horizontal axis. While there is a high probability for holding times of less than 2 hours the logarithmic presentation reveals that there is a small but not negligible probability for long sessions of 20 hours and more. This so-called „heavy tail“ is an indication for high variability of large values ([2]). The coefficient of variation of the holding time data is around 2.8 (i.e. the standard deviation is of the magnitude of 2.8 times the mean value).

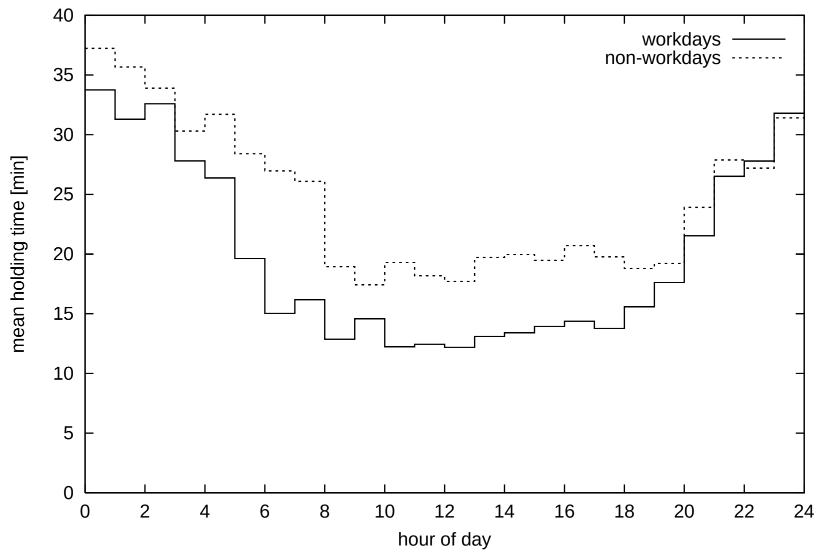

Figure 1: Mean holding time during the day

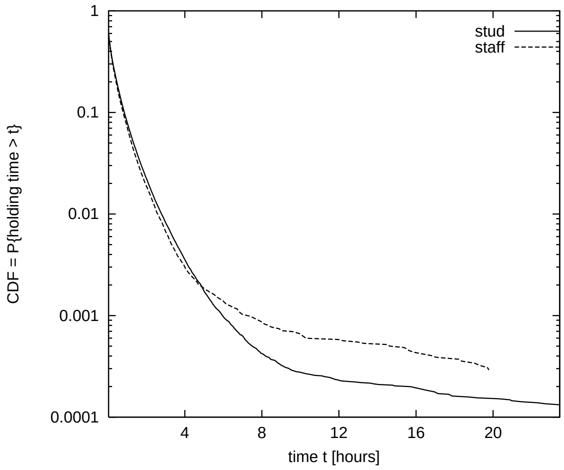

Figure 2: Complementary cumulative distribution function of the holding time

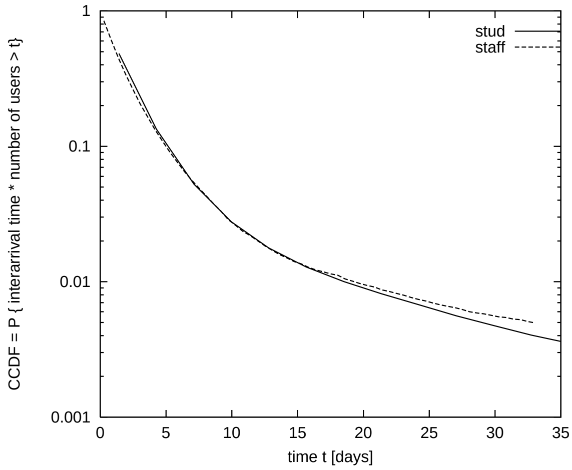

Figure 3: Complementary cumulative distribution function of the interarrival time

2.2 Interarrival time

The interarrival time is the time between two consecutive session beginnings. This time can be measured either for sessions of individual users or for the sessions of all users. The first case leads to a description of the login frequency of a user while the latter one allows a description of the aggregate session arrivals as seen by the access provider.

For the aggregate traffic of 4900 subscribers the mean session interarrival time was 40 seconds. In contrast to the holding time discussed above, the interarrival time of aggregate sessions depends strongly on the number of participating users. The CCDF of the interarrival time of the aggregate traffic would be only a description of the behaviour of a group of 4900 users. To allow a comparison with other data we have scaled this CCDF to describe the interarrival time per user (Figure 3). The scaled curves for 4100 students and 800 staff members fit amazingly well. Again a high variability is found for longer interarrival times which is indicated by the tail of the CCDF.

Figure 4 shows the mean session arrival rate per user during the course of the day for a typical weekday and on weekends. It is obvious that most sessions start during the afternoon and evening. In the early morning far less sessions are originated.

If the interarrival time is measured between sessions of individual users, a completely different distribution is received. Figure 5 shows the resulting complementary cumulative distribution function (average of all individual interarrival times of students as well as of members of staff). The curve shows characteristic steps in regular distances of 24 hours for both user groups. The same behaviour can also be found when regarding smaller user groups like students of certain fields. It is explained by the preferences of the users who tend to go online at a certain time of the day although not necessarily every day. Many users prefer for example to make use of cheaper telephone tariffs starting at 6 pm and 9 pm. Although this periodic behaviour is mostly due to telephone tariffs it is also caused by personal habits or work times of users.

Obviously there is a strong impact of the natural day and night shift, the working hours, individual preferences and tariffing schemes that leads to a periodic behaviour of end users. This periodicity may lead to unfortunate aggregation of traffic load but it may also be useful in influencing user behaviour in order to shift busy hours and balance network load.

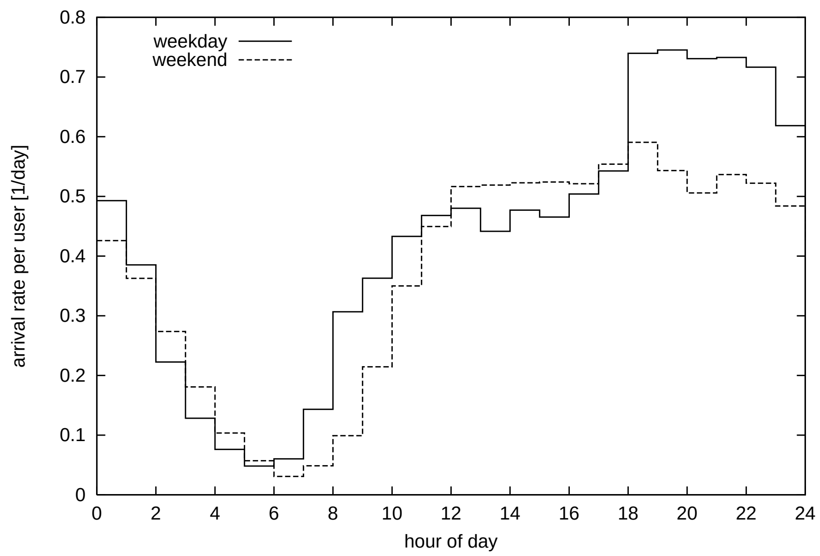

Figure 4: Mean arrival rate per user during the day

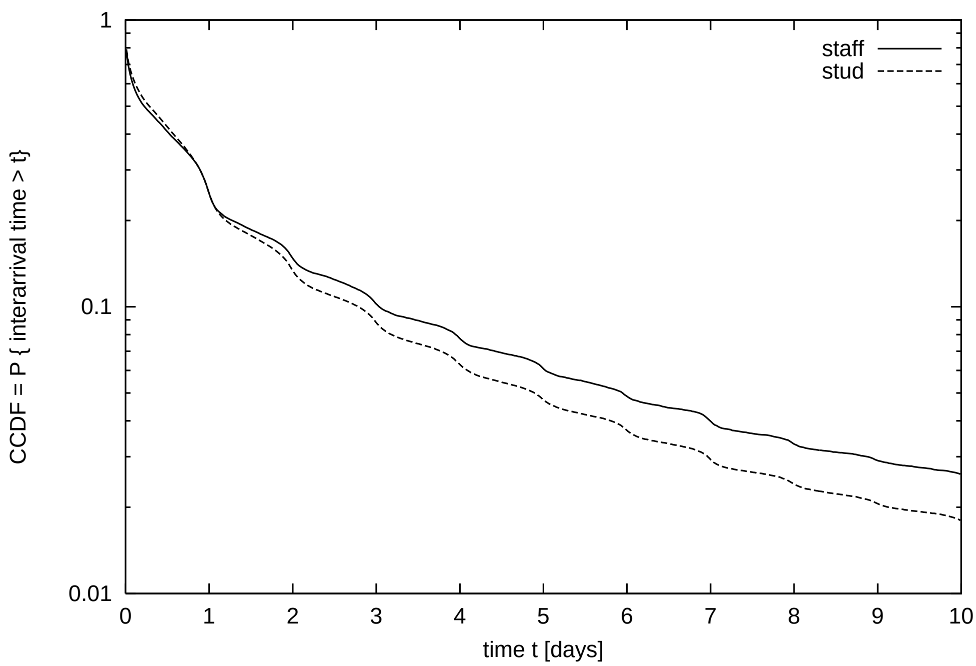

Figure 5: Complementary cumulative distribution function of the session interarrival time for individual users

2.3 Traffic load

To cope with the originated traffic, a telephone network must offer sufficient resources in terms of bandwidth (i.e. telephone lines) and connection setup capacity (i.e. processing power). Therefore we distinguish between traffic load related to user traffic (the seizure of the modem lines) and signalling load corresponding to the call arrival rate (see also Figure 4) to describe the actual load of the access network.

The average values for user traffic load and signalling traffic load for all days of the observed period are shown in the mean daily traffic profile in Figure 6. To allow a better comparison, both values are depicted in the same figure and are drawn in relation to their maximum values (the actual maximum values are 49 Erlang for the mean traffic load and 130 calls per hour for the mean arrival rate - note that only successful calls and no rejected calls could be

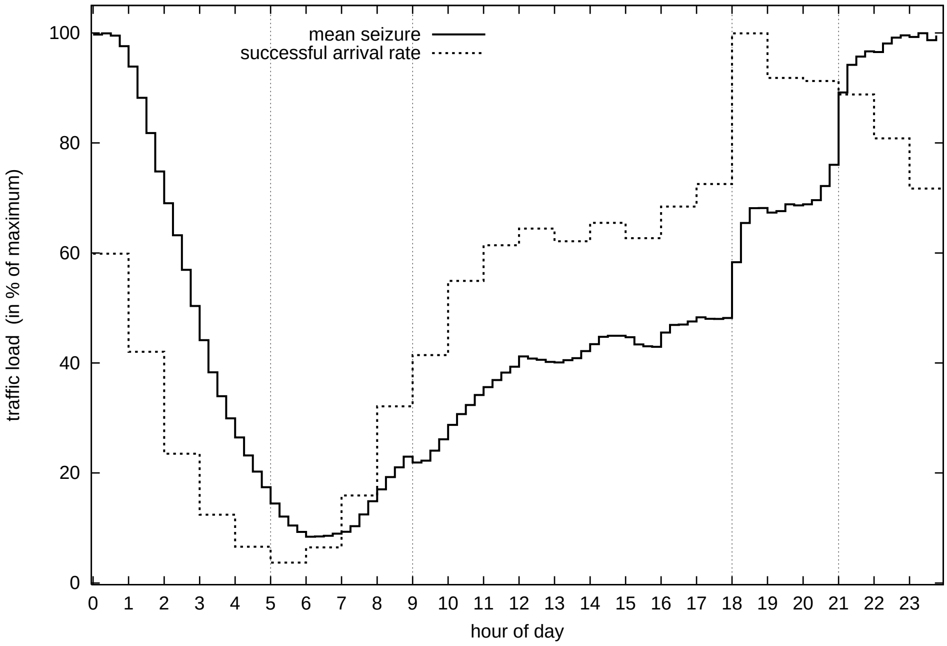

Figure 6: Mean daily traffic profile for traffic load and arrival rate (students only)

detected at the modem pool). For this evaluation only the student modem pool with 56 modems and 4100 subscribers is taken into account.

The general shape of the traffic profile is almost complementary to the classic telephone traffic profile which has its busiest hours during the day and decreases at the end of the day. The shown traffic profile is certainly not typical for dialup sessions in general, but it is typical for the behaviour of home users who usually dial up after work hours.

The most obvious characteristics of the mean seizure are the two steps at 6 pm and 9 pm. As already mentioned above, these times mark the beginning of cheaper telephone tariffs during the observed period. A small peak is also visible right before 9 am when the expensive day tariff starts. On weekends and holidays the day tariff is the same as in the evening and morning hours and therefore no such steps are visible at 9 am and 6 pm. The user behaviour follows these tariffing scheme amazingly accurate.

The flat shape around midnight is due to the limited number of modems. It is interesting that there is a lot of traffic during late night and early morning hours. A comparison with the arrival rate indicates that this traffic is caused by few but long sessions (see Figure 1).

The time consistent busy hour for user traffic load is found at 11:45 pm to 0:45 am and for the arrival rate it is found at 6 pm. In [1] Bolotin points out, that these two busy hours are significantly shifted against each other for Internet traffic compared to telephone traffic. This phenomenon is caused by much longer holding times of around 20 minutes (compared to 3 minutes for classical telephone calls). For design and dimensioning of network components, traffic load has to be evaluated carefully for both busy hours.

3 Modelling dialup session behaviour

For performance evaluation of communication systems by performance analysis or simulation, source traffic has to be modelled to assess the system behaviour. Complex traffic is best described with the help of empirical data. Either a logged traffic trace is replayed into the system model or a mathematical description can be found to generate stochastic traffic of similar characteristics.

To describe traffic load on session level it is important to know about the holding time and the session interarrival time. The complementary cumulative distribution functions of these measures capture their most important characteristics. If mathematical functions can be found that describe the CCDFs, they can be used to parameterize a random generator. This generator may then provide a simulation with values for the measure of interest according to that CCDF.

While classic telephone traffic is well described with negative exponentially distributed holding times and call interarrival times this is not true for Internet access session traffic any more. The high variability of the measures described above is not captured by this distribution. For the description of Internet traffic other distributions have been proposed like Pareto, hyper-exponential or Weibull distributions (e.g. [3], [4], [5], [7]).

In the following sections these distributions are fitted to the empirical distributions of the traces and the resulting parameters are presented (see Table 1). We use the Marquardt-Levenberg algorithm to find a set of parameters for each function. This iterative algorithm performs a non linear least square fit.

Table 1: Cumulative distribution functions of Pareto, Weibull and the hyperexponential distributions

Pareto distribution	Weibull distribution	hyperexponential distribution (order k)
$P\{X \le x\} = 1 - \left(\frac{k}{x}\right)^{\alpha}$ with $k \ge x$	$P\{X \le x\} = 1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$	$P\{X \le x\} = 1 - \sum_{i=1}^{k} p_i e^{-\lambda_i x}$ with $\sum_{i=1}^{k} p_i = 1$

3.1 Holding time

Figure 7 shows the results of the fitting operations to the CCDF of the session holding time. Obviously the Pareto distribution is not suited to describe this measure. The Weibull distribution, on the other hand, leads to a rather good fit and also the hyperexponential distribution describes the behaviour accurately to a certain point. The resulting parameters of the fit can be found in Table 2.

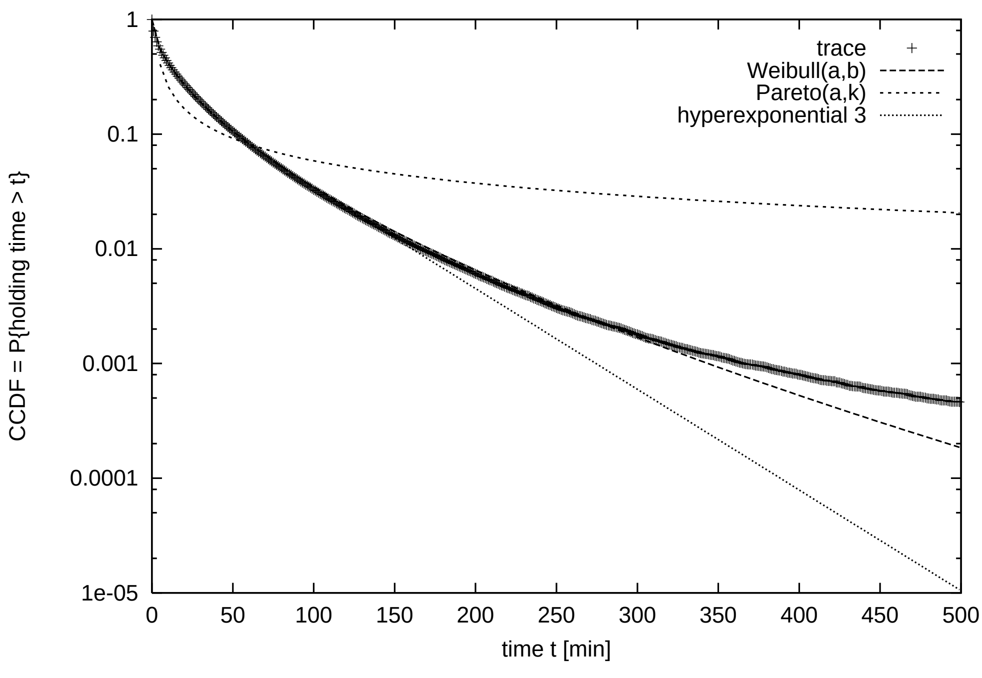

Figure 7: Fitting of several functions to the CCDF of the holding time

Table 2: Resulting parameters of the fitting operation for the holding time

distribution	parameters	mean [min]	CoV
Trace	-	20.6	2.8
Pareto	$\alpha = 0.649$, $k = 1.265$	∞	∞
Weibull	$\alpha = 0.584$, $\beta = 12.544$	19.569	1.82
hyperexponential	$p_1 = 0.436$, $p_2 = 0.257$, $\lambda_1 = 0.072$, $\lambda_2 = 0.020$, $\lambda_3 = 0.73$	19.201	1.71

The table also shows the resulting mean value and the coefficient of variation of the parameterized functions in case they do exist. While the mean values for the Weibull and hyperexponential functions are close to the empirical mean, the variability of the real measure is still significantly higher.

3.2 Interarrival time

As depicted in Figure 8 the scaled CCDF of the interarrival time (of the summary traffic of all subscribers) is best approximated by a Pareto or a hyperexponential distribution. The Weibull distribution can be fitted either to head or tail of the empirical distribution and is not well suited for its description. Table 3 shows the resulting parameters for all fitted functions as well as the corresponding mean values and the coefficients of variation.

Table 3: Resulting parameters of the fitting operation for the interarrival time

distribution	parameters	mean [days]	CoV
Trace	-	2.424	2.4
Pareto	$\alpha = 1.453$, $k = 0.964$	3.095	∞
Weibull	$\alpha = 0.879$, $\beta = 2.091$	2.230	1.272
hyperexponential	$p_1 = 0.926$, $p_2 = 0.064$, $\lambda_1 = 0.528$, $\lambda_2 = 0.137$, $\lambda_3 = 0.031$	2.524	2.040

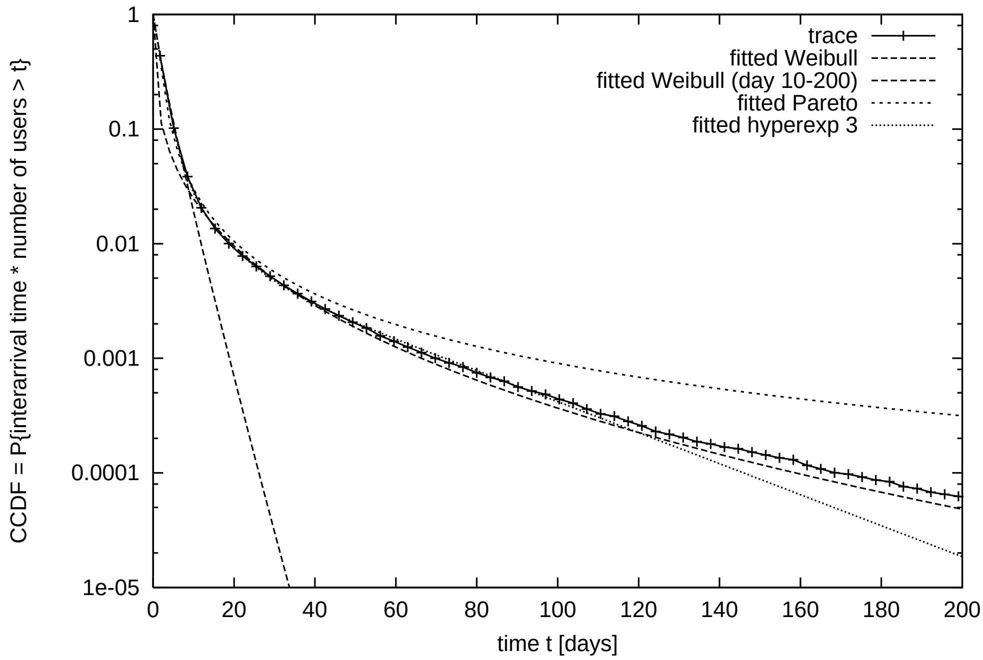

Figure 8: Fitting of several functions to the CCDF of the interarrival time

For the interarrival time as for the holding time, the hyperexponential distribution is difficult to fit to the CCDF because of its numerous parameters. The choice of the starting values is significant and the characteristic heavy tail is not captured accurately. If a hyperexponential distribution is chosen to describe the traffic, a more systematic approach would be preferable as is suggested by Feldmann and Whitt in [4].

4 Conclusion

We have presented the traffic characteristics of dialup sessions monitored at the modem pool of the University of Stuttgart. The long holding times and the high variability of holding time and interarrival time can be found in other publications as well and seem to be typical for Internet traffic. We have also shown that the user behaviour is heavily influenced by the employed telephone tariffing scheme.

Finally a simple mathematical description of the holding time as well as of the interarrival time was presented. Note that the current modelling approach captures the overall behaviour of the traffic. As depicted in Figure 6 the traffic characteristics vary during the course of the day. Therefore, a model describing traffic load only during the busy hours of internet traffic as well as of telephone traffic would be helpful for network dimensioning.

We like to point out, that our results are based on empirical data of a special user group (students and university staff members) and might not describe general Internet traffic. Also the behaviour was strongly influenced by the telephone tariffing scheme in Germany and it should be mentioned that the fast Internet access itself was provided for free.

References

[1] Bolotin V.A.: „New Subscriber Traffic Variability Patterns for Network Traffic Engineering“, Proceedings of the 15th International Teletraffic Congress - ITC 15, Vol 2, pp. 867-878, 1997

[2] Crovella M., Bestavros A.: „Performance Characteristics of World Wide Web Information Systems“, Tutorial at the SIGMETRICS‘97, 1997

[3] Crovella M., Bestavros A.: „Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes“, Proceedings of the ACM SIGMETRICS, pp. 160-169, 1996

[4] Feldmann A., Whitt, W.: „Fitting mixtures of exponentials to long-tailed distributions to analyze network performance models“, Performance Evaluation 31, pp. 245-179, 1998

[5] Mah, B.A.: „An Empirical Model of HTTP Network Traffic“, Proceedings of the IEEE INFOCOM´97, Vol 2, pp. 592-600, 1997

[6] Morgan, S.: „The Internet and the Local Telephone Network: Conflicts and Opportunities“, IEEE Communications Magazine, pp. 42-48, January 1998

[7] Shuang D.: „Empirical Model of WWW Document Arrivals at Access Link“, ICC International Communication Conference, 1996