MODELLING TELEPHONE ACCESS TO
INTERNET SERVICE PROVIDERS

Dimitrijević, D.
Consultant

I INTRODUCTION

In the old days of Plain Old Telephone Service (POTS), Erlang-B formula [1, 2] was the way to estimate the number of telephone lines necessary to handle the offered traffic. Standard assumptions were: (a) independent and exponentially distributed times between calls, and (b) independent and exponentially distributed duration of calls with mean value of 3 minutes. With the introduction of various Internet applications, the nature of telephone calls has been significantly changed. In this paper we examine how to model telephone access to Internet Service Providers (ISP).

II PROBLEM

Modelling of usage of telephone lines in the POTS assumes independent and exponentially distributed interarrival and service times. It is assumed that the incoming traffic is homogeneous, i.e., all calls have the same nature (voice calls). This is not true in the case of access to ISPs, since the purpose of calls may be different. For example, one may access the ISP to retrieve/send e-mail. In this case, duration of a call is relatively short. In the other case, one may hold the line busy during a long session while surfing the Internet. Each type of calls may have its own characteristics (arrival rate and duration of calls). These parameters may be influenced by ISP's billing strategy. For example, some ISPs have "lite" billing rate (cheaper) which allows users to check e-mail and access ISP's web site. To surf the entire Internet, users must log-on using "full" rate.

To illustrate various types of traffic, we use author's call records during one month period. Since the author is not a typical user, this data should be used as an illustration of the concept, not to make any quantitative conclusions. Figure 1 shows the Cumulative Distribution Function (CDF) of call duration of short calls made to retrieve/send e-mail. Figure 2 shows the CDF of call duration of long calls made to surf the Internet.

There have been 573 short and 80 long calls. Their average duration was 33 and 337 seconds respectively. Certainly, these numbers depend very much on a particular user's behavior and characteristics of a particular computer (speed of processor, disk, and modem). In this case, the two CDFs do not have exponential behavior. However, a large number of users would result in a larger variance of call duration.

The above mentioned parameters may influence ISP's revenue. Even in the case of a flat monthly rate, the ISP must insure a satisfactory quality of service (low probability of busy lines), and have a sufficient number of telephone lines (often in the blocks of 120 lines). In the following section we develop a multiple class Erlang-B formula to handle multiple streams of various types of traffic.
III SOLUTION

Let there are \( N \) input telephone lines to the ISP system. Let us assume that \( K \) independent streams of calls are coming to the system. Interrival and service times of each stream \( k \) are independent and exponentially distributed with rates \( \lambda_k \) and \( \mu_k \). The non-negative number of active calls of each stream \( k \) is shown in Figure 3, of each stream \( k \) define the state of the system \( \left(i_1, i_2, \ldots, i_k\right) \) . The state space is bounded by the constraint that the total number of active calls is not greater than the number of telephone lines, i.e., \( \sum_{k=1}^{K} i_k \leq N \) . For simplicity of the drawing, a two-dimensional state transition diagram, for two input streams, is shown in Figure 3. The state space has a triangular shape. Only transitions in and out of one inner state are shown.

![Figure 3: Two-dimensional state transition diagram](image)

Let \( p_{i_1, i_2, \ldots, i_k} \) be the probability that the system is in state \( \left(i_1, i_2, \ldots, i_k\right) \) . To handle the boundary cases in a uniform way, we introduce function \( h(x) \). It is equal to 1 when \( x \geq 0 \) and 0 otherwise. Then, the balance of flow equation around state \( \left(i_1, i_2, \ldots, i_k\right) \) may be written as

\[
\begin{align*}
&h\left(N-1 - \sum_{k=1}^{K} i_k \right) \sum_{k=1}^{K} \lambda_k p_{i_1, i_2, \ldots, i_k} + \sum_{k=1}^{K} \lambda_k p_{i_1, \ldots, i_k-1, \ldots, i_k} + \\
&\left(N-1 - \sum_{k=1}^{K} i_k \right) \sum_{k=1}^{K} \lambda_k p_{i_1, \ldots, i_k-1, i_k} + \sum_{k=1}^{K} (i_k + 1) \mu_k p_{i_1, \ldots, i_k+1, \ldots, i_k} + 1. 
\end{align*}
\]

By replacement, it is easy to verify that

\[
P_{i_1, i_2, \ldots, i_k} = S_N \prod_{k=1}^{K} \left( \frac{\lambda_k}{\mu_k} \right)^{i_k}
\]

satisfies the system of linear equations (1). The \( S_N \) is the normalizing value chosen in such a way that the sum of all probabilities is equal to 1. Let us define \( \rho \) as

\[
\rho = \sum_{k=1}^{K} \frac{\lambda_k}{\mu_k}
\]

Then \( S_N \) can be calculated as

\[
S_N^{-1} = \sum_{n=0}^{N} \frac{\rho^n}{n!} = \sum_{n=0}^{N} \frac{\rho^n}{n!} = \sum_{n=0}^{N} \frac{\rho^n}{n!}
\]

The blocking probability, i.e., the probability that all lines are busy, can be calculated using the Erlang B formula

\[
P_b(N) = \frac{S_N}{N!} \sum_{i_1+\ldots+i_K=N} N! \prod_{k=1}^{K} \frac{1}{i_k!} \left( \frac{\lambda_k}{\mu_k} \right)^{i_k} = \frac{p_b(N-1)}{N!} \sum_{n=0}^{N} \frac{\rho^n}{n!}
\]

Next, we define a procedure to find the most likely state of the system, i.e., the state that maximizes probability (2). The procedure runs as follows:

1. Set \( i_k \leftarrow 0 \) for all \( k \).
2. If \( \sum_{k=1}^{K} i_k = N \) then stop.
3. Find \( k \) such that \( A = \frac{\lambda_k}{\mu_k} \) is maximal. If \( A \leq 1 \) then stop.
4. For the \( k \) found in step 3 set \( i_k \leftarrow i_k + 1 \) and go back to step 2.

When the procedure stops, \( \left(i_1, i_2, \ldots, i_k\right) \) defines the most likely state. Physically, it means that the system spends most of the time in that state and the surrounding states in the state transition diagram. Computational complexity of a straightforward implementation of the described procedure is \( O(NK) \) . A more careful implementation would use a data structure called heap [3], to execute step 3 more efficiently. In that case, computational complexity is \( O(N \log K) \) . Such an approach was used in [4] to optimize a similar function with discrete valued parameters.
Finally, let us note that the described system is analytically equivalent to a single stream of independent and exponentially distributed arrivals with the arrival rate

\[ \Lambda = \sum_{k=1}^{K} \lambda_k \]  

(6)

and the hyper-exponentially distributed service time with the probability density function defined as

\[ f(t) = \sum_{k=1}^{K} \frac{\lambda_k}{\Lambda} \mu_k e^{-\mu_k t} \]  

(7)

IV CONCLUSIONS

In this paper we developed a procedure for modelling telephone access to Internet Service Providers. This is one of the steps toward insuring satisfactory performance of the entire ISP's system.

REFERENCES


Abstract: In this paper we model telephone access to Internet Service Providers (ISP). Unlike voice calls which represent a single class of calls, the developed model assumes multiple classes of calls such as retrieval of e-mail messages or surfing the Internet. The blocking probability, i.e., the probability that all telephone lines are busy, is a generalization of the Erlang-B formula. We also developed a procedure to determine the most likely state of the system.

MODELLING TELEPHONE ACCESS TO INTERNET SERVICE PROVIDERS, Dragomir D. Dimitrijević.