THE EFFECT OF HOLDING TIMES ON LOSS SYSTEMS

V. B. IVERSEN
Technical University of Denmark, Lyngby, Denmark

ABSTRACT
This paper considers loss systems with Poissonian arrival processes and general holding time distributions.

It is well-known that the classical formula for full available groups of Erlang, Engset and Palm (machine/repair, one repairman) are valid for arbitrary holding time distributions.

For limited available groups the probability of loss depends in general on the type of holding time distribution. But it is shown that Erlang's Interconnection Formula is valid for general holding time distributions.

Quantitative investigations of general systems with limited availability show that (as opposed to delay systems) constant holding times yield maximum probability of loss. If the form factor of the holding time increases, then the probability of loss decreases.

The results are of importance, when evaluating previous works based on the assumption of exponential holding times, and when studying the optimal structure of gradings and link systems.

INTRODUCTION
Loss systems have in general only been studied for pure chance traffic. This is mainly due to the limited tools available for traffic engineers: birth & death processes and roulette simulation. A multitude of investigations into loss systems have been based on these assumptions: Poisson arrival process (possibly with a limited number of sources) and exponential holding times. In this model the total traffic process is described by only one parameter, viz. the mean value of the offered traffic.

The successful application of teletraffic theory demonstrates that these assumptions in general are sufficient.

Observations of real traffic (Iversen 1973) shows that the call arrival process is a Poisson process having slow intensity variations during the day.

However, the holding times are seldom exponential distributed. But the success of teletraffic theory is due to the fact that loss systems are independent or almost independent of the holding time distribution.

In the following we assume Poisson arrival process and general holding time distributions.

Very few practical investigations have been published on the sensitivity of the classical assumptions to deviations from the exponential holding time distribution, except for the classical full available models (ref. e.g. Wolff 1977).

Some highly theoretical studies have dealt with these problems for Semi-Markov processes, e.g. König, Matthes & Nawrotzki (1971), and Schassberger (1977, 1978). However, these studies are difficult to comprehend for traffic engineers, and the application of the results to traffic engineering has not yet been demonstrated.

TIME DISTRIBUTIONS
Every holding time is a non-negative stochastic variable $T$, which is characterized by a lifetime distribution function $F(t)$, or by a density function $f(t)$:

$$ F(t) = \begin{cases} 1 & t > 0 \\ 0 & t < 0 \end{cases} \\ f(t) = \frac{dF(t)}{dt} $$

A distribution is determined by its moments $M_n$ defined by:

$$ M_n = \int_0^\infty t^n f(t) \, dt $$

The most important moment is the mean holding time $s$:

$$ s = M_1 $$

In the following we shall always assume that the mean holding time exists, even when we talk about general distributions.

The second moment is usually transformed to the variance $\sigma^2$ or to the form factor $\varepsilon$:

$$ \varepsilon = \frac{M_2}{s^2} = 1 + \frac{\sigma^2}{s^2} $$

From observations we are usually only able to obtain relevant information about the two first moments, and therefore higher order moments play a minor role in teletraffic engineering.

When calls arrive at a full available group according to a Poisson process, the mean holding time yields full information for the calculation of the congestion. In other cases we also need to know the form factor. The variance of a measurement of a traffic intensity is thus given by:

$$ \sigma_1^2 = \frac{s^2}{A} \cdot \varepsilon $$

where $A$ is the mean traffic intensity, $T$ the measuring period, and $\varepsilon$ the form factor of the holding times. We obtain the smallest variance for constant holding times ($\varepsilon=1$). For real
The exponential distribution:

$$f(t) = u \cdot e^{-ut} \quad u > 0, \quad t > 0$$

has the mean value $s = 1$ and the form factor $\varepsilon = 2$.

Erlang-k distribution

Time distributions having a form factor $1 < \varepsilon < 2$ are called steep distribution functions.

They can be obtained by putting exponential distributions in series. Every exponential term is then called a phase. If all phases have the same intensity, then the form factor becomes

$$\varepsilon = 1 + \frac{1}{k}$$

where $k$ is the number of phases. A simple example is shown in fig. 1.

![Fig. 1: Erlang-2 distribution having mean value $1/u$ and form factor 1.5.](image)

The distribution shown in fig. 1 has the density function:

$$f(t) = 4u^2 t \cdot e^{-2ut}$$

If we denote the arrival intensity by $\lambda$, then the total offered traffic is

$$A = \lambda / \mu$$

Each of the two phases contributes with a traffic:

$$A_1 = A_2 = A/2$$

Hyper-exponential distribution

Upon the other hand we can put exponential terms in parallel and obtain a distribution, which has a form factor $2 < \varepsilon < \infty$.

This is called a flat distribution, and a well-known example is shown in fig. 2.

![Fig. 2: Hyper-exponential-2 distribution.](image)

If we put $p_1 = p = 1/10$, $p_2 = 1-p = 9/10$, $\mu_1 = 1/7$, and $\mu_2 = 3$, we get the density function:

$$f(t) = \frac{1}{70} e^{-t/7} + \frac{9}{10} \cdot 3 \cdot e^{-3t}$$

This distribution has the mean value $s = 1$ and the form factor $\varepsilon = 10$. The phases contribute with the following traffic parts:

$$A_1 = \lambda \cdot \frac{p}{\mu_1} = 0.7 A$$

$$A_2 = \lambda \cdot \frac{1-p}{\mu_2} = 0.3 A$$

Coxian distribution

Above we have mentioned 3 typical distributions having form factor 1.5, 2, resp. 10.

In principle we may combine exponential phases in parallel/serial and obtain generalized Erlang-distributions as approximations to a general distribution. This is done by means of a Coxian distribution (Bux & Herzog 1977) shown in fig. 3.

This distribution has a rational Laplace transform. Since any Laplace transform can be approximated arbitrarily closely by rational functions, the Coxian distribution can be approximated arbitrarily closely to any distribution.

If we are able to show that a formula is valid for Coxian distributions, then it is valid for a general distribution. This we shall use in connection with Erlang's Interconnection Formula.

The mean value of the Coxian distribution shown in fig. 3 is:

$$s = \frac{r}{\sum_{i=1}^{r} \frac{q_i}{\nu_i}}$$

where

$$q_i = p_0 \cdot p_1 \cdot \ldots \cdot p_{i-1}$$

The total offered traffic is

$$A = \lambda \cdot s = \frac{r}{\sum_{i=1}^{r} A_i}$$

where the traffic offered from phase $i$ is

$$A_i = \lambda \cdot \frac{q_i}{\nu_i}$$

Whereas the Erlang-2 distribution is a Coxian distribution, the hyper-exponential distribution is not. However, it is easy in a formula to replace a Coxian distribution by a hyper-exponential distribution, as we shall see later.

The form factor of the Coxian distribution may take any value

$$1 < \varepsilon < \infty$$

The above family of generalized Erlang distributions represents such a wide range of distributions that there is no reason to use other time distributions.
ERLANG’S INTERCONNECTION FORMULA

A.K. Erlang (1920) gave an exact solution for the blocking probability of an ideal grading having $n$ lines and the availability $k$. The number of inlet groups $g$ is determined from $n$ and $k$:

random hunting: $g_{rh} = \binom{n}{k}$
sequential hunting: $g_{sh} = \binom{n}{k} \cdot k!$

A mathematical deduction of the Interconnection Formula was given by E. Brockmeyer (1948), who also has pointed out the assumptions:

a. A homogeneous group of $n$ lines is considered.
b. The offered traffic is Pure Chance Traffic type I (Poissonian arrival process with intensity $\lambda$, and exponential holding times with mean value $1/\mu$).
c. Every call attempt has access to $k$ lines.
d. The call attempts hunt the $k$ trunks in such a way that the busy lines at every point of time are distributed at random among the $n$ lines.
e. A call attempt, which finds all $k$ lines busy, is lost.

Under these assumptions the number of states is reduced to $n+1$ different states as in the case of full availability, and we are able to calculate the state probabilities and the blocking probability. The probability of observing $v$ lines busy is denoted by $[v]$, and it is readily shown that

$$[v] = \frac{Q_v \cdot A^v}{\sum_{v=0}^{n} Q_v \cdot A^v}, \quad v=0,1,...,n$$

where

$$Q_v = \frac{n!}{v! \cdot (n-v)!} \cdot (1-B_{v/})$$

and

$$B_v = \binom{n}{k} \cdot \begin{cases} 0 & 0<v<k \\ \frac{v(v-1)...(v-k+1)}{n(n-1)...(n-k+1)} & k<v<n \end{cases}$$

The probability of blocking is

$$E = \sum_{v=0}^{n} B_v \cdot [v]$$

For $k=n$ we obtain Erlang’s B-formula as a special case.

Erlang’s Interconnection Formula represents, generally speaking, the lower limit for the congestion in a grading. The formula has not yet obtained the deserved application. This is due to the number of parameters ($n, k, A$) and to the complexity of working it out. However, by means of a programmable calculator this is very easy (Iversen 1977).

ARBITRARY HOLDING TIMES

We now give the state probabilities for Erlang’s ideal grading (Interconnection Formula) having Coxian holding times.

Let

$$\{ x = x_1 + x_2 + ... + x_r \}$$

denote the probability that $x_i$ lines are in phase $i$ ($i=1,r$) and that totally $x$ lines are busy.

The marginal probability $[x]$, that totally $x$ lines are busy, becomes the same as for Erlang’s Interconnection Formula. Thus this probability of observing "x lines busy" shall be split up into individual micro states, and this must be done according to the following rule (cf. the polynomial distribution):

$$\{ x = x_1 + x_2 + ... + x_r \} = \binom{x}{x_1, x_2,..., x_r} \cdot \frac{A_{x_1}^1 \cdot A_{x_2}^2 \cdots A_{x_r}^r}{A^x} \cdot [x]$$

where

$$\binom{x}{x_1, x_2,..., x_r} = \frac{x!}{x_1! \cdot x_2! \cdots x_r!}$$

and $A_i$ is the traffic caused by phase $i$ (ref. holding times)

The state $\{ x = x_1 + x_2 + ... + x_r \}$ is itself a macro state consisting of

$$\binom{n}{x_1, x_2,..., x_r} \cdot \frac{1}{(n-x)!}$$

equally probable (because of symmetry) micro states.

Thus one micro state belonging to the set $\{ x = x_1 + x_2 + ... + x_r \}$ has the state probability

$$\{ x = x_1 + x_2 + ... + x_r \} = \frac{A_{x_1}^1 \cdot A_{x_2}^2 \cdots A_{x_r}^r}{A^x} \cdot [x]$$

This result is intuitively correct and very simple.

The formula is in the general case difficult to prove from the state diagram, but we shall exemplify it by two examples: Erlang-2 distributed holding times and hyper-exponential-2 distributed holding times.

NUMERICAL EXAMPLES

In both cases we choose $n=3$ lines, $k=2$, and $A=2$ erlang. This grading is shown in fig. 4

$$\lambda \quad \frac{1}{3} \quad \lambda \quad \frac{1}{3} \quad \lambda \quad \frac{1}{3} \quad 2 \quad 3$$

Fig. 4: Erlang’s ideal grading ($n,k$) = (3,2) for random hunting.

From Erlang’s Interconnection Formula we find the marginal probabilities:
We choose the above-mentioned Erlang-2 distribution, where each phase has the intensity $2u$. As we have a high degree of symmetry:

$$A_1 = A_2 = A = \frac{\lambda}{2u}$$

we have a high degree of symmetry:

$$[0] = \frac{1}{1 + A + \frac{A^2}{2T} + \frac{2 A^3}{3T^2}} = \frac{9}{53}$$

$$[1] = A \cdot [0] = \frac{18}{53}$$

$$[2] = \frac{A^2}{2T} \cdot [0] = \frac{18}{53}$$

$$[3] = \frac{2}{3} \cdot \frac{A^3}{3T^2} \cdot [0] = \frac{5}{11}$$

$$E = \left( \frac{1}{2} \frac{A^2}{2T} + \frac{2}{3} \cdot \frac{A^3}{3T^2} \right) \cdot [0] = \frac{14}{53}$$

### ERLANG-2 HOLDING TIMES

We choose the above-mentioned Erlang-2 distribution, where each phase has the intensity $2u$.

$$A_1 = A_2 = A = \frac{\lambda}{2u}$$

we have a high degree of symmetry:

$$[0] = \frac{1}{1 + A + \frac{A^2}{2T} + \frac{2 A^3}{3T^2}} = \frac{9}{53}$$

$$[1] = A \cdot [0] = \frac{18}{53}$$

$$[2] = \frac{A^2}{2T} \cdot [0] = \frac{18}{53}$$

$$[3] = \frac{2}{3} \cdot \frac{A^3}{3T^2} \cdot [0] = \frac{5}{11}$$

$$E = \left( \frac{1}{2} \frac{A^2}{2T} + \frac{2}{3} \cdot \frac{A^3}{3T^2} \right) \cdot [0] = \frac{14}{53}$$

### CONSTANT HOLDING TIMES

It may also be mentioned, that a time-true simulation with constant holding times comprising 5 million call attempts came out with the following result:

- mean congestion = 0.26405
- standard deviation = 0.00018

The theoretical congestion is

$$E = \frac{14}{53} = 0.26415$$

### GRADINGS IN GENERAL

Above we have shown that Erlang's Interconnection Formula is independent of the holding time distribution. As a special case $[k=n]$ we obtain Erlang's loss formula. We may ask whether this is true for any grading. The answer is no.

The grading shown in fig. 5 is so simple that we easily can obtain the blocking probability for the previous mentioned 2-phase distributions. For $A=2$ erlang we find analytically:

- Erlang-2: $E_1 = 1.5$  $E_2 = 2$
- Exponential: $E_1 = 0.454614$  $E_2 = 0.454545$
- Hyperexponential: $E_1 = 0.454379$

These results are typical for loss systems with Poissonian arrival processes:

a. The dependence on the form factor (i.e. the distribution function) is very small. It is e.g. difficult to demonstrate by numerical simulation.

b. The worst case is constant holding times ($\epsilon = 1$).

In loss systems it is an advantage with a large form factor. This is diametrically opposed to delay systems, where the waiting time in general is proportional with the form factor.

This result is in agreement with R.W. Wolff (1977) and Nilsson (1978). A more extensive treatment of this problem is given in (Iversen & Nielsen 1979).

### CONCLUSION

We are able to draw several important conclusions, which are of importance for practical teletraffic engineering:

- Systems with limited availability.
- Erlang's ideal grading is independent of the holding time distribution and thus more fundamental than considered in general.
The more symmetric a grading is, and the more uniform the load is distributed on the lines, the more independent is the grading of the holding time distribution. Thus a symmetrically loaded ideal grading with random hunting is almost independent of the holding time distribution. More details are given in (Iversen & Nielsen 1979).

Simulation

Simulation of loss systems may be done by the roulette method or by true-time simulation. The roulette method is based on exponential holding times. When simulating in true time we should always use constant holding times as this has several advantages:

a. This is the worst case (maximum congestion).

b. The variance of the traffic load and other parameters are in general proportional to the form factor of the holding time. Using constant holding times thus allows us to reduce the length of the simulation with a factor $\sqrt{2}$ as compared with exponential holding times.

c. The generation of constant holding times is fast and accurate.

Traffic measurements

If the mean holding time is known, we should not measure the traffic volume, but count the number of calls. This, of course, presumes the mean holding time is independent of the arrival process.

Further research

It would be extremely useful to have a theorem stating that a system, which - for a given load - is optimal for one value of the form factor, is also optimal for any other value of the form factor.

Final remarks

The success of the classical teletraffic theory is due to the facts that loss systems in general are almost independent of the holding time distribution, and that the arrival process is a Poisson process (Iversen 1973). Optimal Systems (full availability, Erlang’s ideal grading) are independent of the holding time distribution.

Most studies in this field are of a very theoretical nature and not useful for practical applications. Above we have presented the basic state of things, which are of practical importance for traffic engineering of loss system.

REFERENCES


