Erlang B formula and its corresponding Markov Chain

Sunday, November 29, 2009
9:36 PM

Erlang B formula

In this note, we return to the Erlang B formula.

$$P_b = \frac{\frac{A^m}{m!}}{\sum_{i=0}^{m} \frac{A^i}{i!}}$$

call blocking probability

A = traffic intensity $= \frac{\lambda}{\mu}$

λ = average # of call attempts/requests per unit time

This λ is the rate for the entire trunked system. (not per user)

$\frac{1}{\mu} = H$ = average call length.

This blocking probability is a measure of the grade of service for a trunked system that provides no queuing for blocked calls.

We will see how we get this formula from a Markov chain.

We will again consider small time intervals.

Recall that one of the assumption we made to get the Erlang B formula is that traffic requests are described by a Poisson process (PP) which implies

(1) exponentially distributed call interarrival time and

(2) independence among interarrival times of call request.

These are facts that we proved when we talked about PP.

By saying that the arrival process is Poisson, I don't need to talk about these properties; they automatically follow.

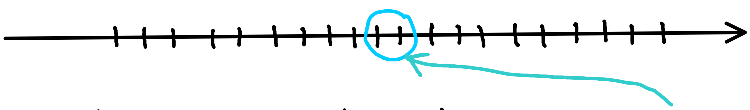
Now, in this system, there are M channels available in the trunking pool. Therefore, the probability that a call requested by a user will be blocked is given by

$$P_b = \underset{\sim}{P}\left[\text{None of the } m \text{ channels are free}\right]$$

this is my notation for "the probability that"

We will consider the long-term behavior of this system, i.e. the system is assumed to have been operating for a long time. In which case, at the instant that somebody is trying to make a call, we don't know how many of the channels are currently free.

To analyze this system, let's first divide the time into small slots all of which occupy the same length (as we have alway done.)



Then, consider a particular slot, say, this one. Suppose that at the begining of this slot, there are K channels that is currently used. We want to find out how this number K changes as we move forward one slot time.

This random variable K will be called the "state" of the system.

This is the same "state" concept that you have studied in digital circuits class.

There are many possible values for the "state" and the system moves from one state to another one as we move forward in time by one slot.

For example, suppose there are 5 persons using the channels at the beginning of the slot. Then K=5. Suppose by the end of the slot, none of these 5 persons finish their calls. Suppose also that there is one more person want to make a call at some moment of time during this slot. Then, at the end of the slot, the number of channels that are used becomes

$$5 - 0 + 1 = 6.$$

So, the state K of the system changes from 5 to 6 when we reach the end of slot which is also the beginning of the next slot.

Again, we want to study how the state K changes from one slot to the next slot.

It might be helpful to label the state K as

$K(1)$ for the first slot

$K(2)$ for the second slot

and so on

so that we know what K you are talking about.

As the example above shown, to know how K changes, we need to know two info:

1) how many calls (that are being made at the beginning of the slot) end during the slot that we are considering?

2) how many new call requests are made during the slot that we are considering?

It turns out that we already know the answer to the second question.

If the interval are small enough (say length $=\delta$) then there can be at most one new arrival (new call request) which occurs with probability

$$p_1 = \lambda \times \delta$$

Of course, you may not trust this approximation. In which case, I'd ask you to go back and read our note on PP.

Alternatively, I want to approach this approximation from another perspective.

When you studied PP for the first time, usually, you would be told that a defining property of a PP is that if you consider non-overlapping intervals of length $T_1, T_2, \dots, T_n$, and then count the number of arrivals $N_1, N_2, \dots, N_n$ in them. Then, all these N_i will be independent

Poisson random variable with mean $\lambda \times T_i$.

Of course, we derived this fact from a more primitive assumption in our note. But let's assume that we start our belief about PP with the above fact, then how can we arrive at the red fact above??

It turns out that this is easy. For an interval of length δ, if we assume Poission r.v., then

$$\underbrace{P[N=n]}_{\text{number of arrivals in interval of length } \delta} = e^{-\lambda\delta}\frac{(\lambda\delta)^n}{n!}$$

In particular

$$P[N=0] = e^{-\lambda\delta} \quad \text{and}$$

$$P[N=1] = \lambda\delta\, e^{-\lambda\delta}$$

Recall, from calculus class, that

$$e^{x} = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots$$

and

$$x e^{-x} = x - x^2 + \frac{x^3}{2} - \frac{x^4}{6} + \cdots$$

all of these are simply Taylor's series around $x=0$.

If you are lazy, you may check these expansions in MATLAB by using

```
syms x
pretty(taylor(exp(x)))
pretty(taylor(x*exp(-x)))
```

Another easy way to find the expansion of xe^{-x} is to multiply the expansion of e^{-x} by x.

When x is small, $x^2, x^3, x^4, \ldots$ are even smaller! So, we will ignore them.

So, we will ignore them.

Hence

$$\left.\begin{aligned} e^{x} &\approx 1+x \\ x e^{-x} &\approx x \end{aligned}\right\} \text{when } x \text{ is small.}$$

Now, back to our Poisson probability:

$$\left.\begin{aligned} P[N=0] &= e^{-\lambda\delta} \approx 1-\lambda\delta \\ P[N=1] &= \lambda\delta\, e^{-\lambda\delta} \approx \lambda\delta \end{aligned}\right\} \text{when } \lambda\delta \text{ is small}$$

(because λ is a fixed constant, this statement is true when δ is small)

Notice that

$$P[N=0] + P[N=1] \approx (1-\lambda\delta) + \lambda\delta = 1!$$

These two events already take all the probability. In other words, when δ is small, the only two events that can occur is $N=1$ or $N=0$, which is what we already learn from the PP note earlier.

Then we have the first question: how many calls finish during a slot?

For this one, we will use another assumption of the Erlang B formula, which is that the probability of a user occupying a channel (called the service time) is exponentially distributed.

Review: Exponential random variable (r.v.) X

* It is a <u>continuous</u> r.v..
* Probability density function (pdf) is given by

$$f_X(x) = \mu e^{-\mu x}, \quad x > 0.$$

* $P[X>x] = \int_x^\infty f(\tau)\,d\tau = \int_x^\infty \mu e^{-\mu\tau}\,d\tau$

$$= -e^{-\mu\tau}\Big|_x^\infty = e^{-\mu x}$$

a.k.a. CCDF

* Cumulative distribution function (cdf) is given by

$$F(x) = P[X \le x] = 1 - P[X > x] = 1 - e^{-\mu x}$$

* Memoryless Property ← The term memoryless

has many uses in probability theory. This is one of them.

$$P[X > \alpha + \delta \mid X > \alpha] = P[X > \delta]$$

To see this, note that

$$P[X > \alpha + \delta \mid X > \alpha] = \frac{P[X > \alpha + \delta \text{ \underline{and} } X > \alpha]}{P[X > \alpha]}$$

However, the event $[X > \alpha + \delta$ and $X > \alpha]$ is the same as the event $[X > \alpha + \delta]$ because the event $[X > \alpha]$ is already a part of $[X > \alpha + \delta]$. When you need $[X > \alpha + \delta]$, it automatically implies that $[X > \alpha]$.

To understand this, let suppose you want the probability that $X > 5$ and $X > 3$. Of course the part that says $X > 3$ needs not be there. By requiring that $X > 5$, it automatically > 3. So

$$P[X > 5 \text{ and } X > 3] = P[X > 5].$$

Similarly,

$$P[X > \alpha + \delta \text{ and } X > \alpha] = P[X > \alpha + \delta].$$

Plugging this back into the conditional probability above, we have

$$P[X > \alpha + \delta \mid X > \alpha] = \frac{P[X > \alpha + \delta]}{P[X > \alpha]} = \frac{e^{-\mu(\alpha + \delta)}}{e^{-\mu\alpha}}$$

$$= e^{-\mu\delta} = P[X > \delta]$$

What does this memoryless property mean? Suppose you have a light bulb and you have used it for 5 years already. You don't know when it will fail to work after today. Let its lifetime be X, then of course X is a random variable. You also know that $X > 5$ because it is still working now. Suppose you want to find out whether it will still be ok after two more years of use, then you want to find

$$P[X > 7 \mid X > 5]$$

If X is an exponential random variable, then

If X is an exponential random variable, then the memoryless property holds and you would get

$$P[X > 7 \mid X > 5] = P[X > 2]$$

In other words, the likelihood that you can use it for at least two more years is exactly the same as the probability that you can use a new lightbulb for more than two years.

So, your old lightbulb essentially forgets that it has been used for 5 years. It always act as a new lightbulb (statistically).

This is what we mean by the memoryless property of an exponential random variable.

For us, to get an answer for the first question of how many persons finish their calls during our small interval of interest, we will use the memoryless property of exponential random variable.

First recall one parameter in the Erlang B formula. We have

$$A = \frac{\lambda}{\mu}$$

where the $\frac{1}{\mu}$ is the average call duration. This definition, along with the assumption that the call duration is exponentially distributed, gives us the exact pdf of the call duration:

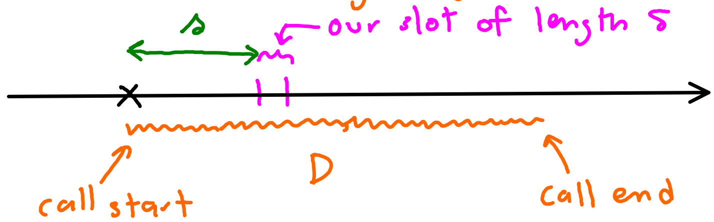
$$f_D(d) = \mu e^{-\mu d}, \quad d > 0.$$

Check : $\displaystyle\int_0^\infty t f_D(t)\,dt = \int_0^\infty t\mu e^{-\mu t}\,dt = \frac{1}{\mu}.$

Now, we can return to our small slot. Let's consider each of the calls that is currently using the channel(s). We know, by memoryless property of exponential r.v. that we don't have to know when these calls started. The remaining duration of them will still be exponential with rate μ. Now, the width of our interval is δ. The probability that a particular call, which is in progress at the beginning of the slot, to be unfinished by the end of the slot is

$$P[D > s+\delta \mid D > s] = P[D > \delta] = e^{-\mu\delta} \approx 1-\mu\delta$$

s = length of time from the time that the call starts to the beginning of our slot



Consequently the probability that a particular call ends during our slot time is $\approx 1-(1-\mu\delta)$
$= \mu\delta$.

Of course, there are many of these calls that are ongoing at the beginning of our slot.

The probability that none of them finishes is

$$\left(e^{-\mu\delta}\right)^{k} \approx 1-k\mu\delta.$$

if $K = k$

The probability that exactly one of them finishes is

$$k\,\mu\delta\left(e^{-\mu\delta}\right)^{k-1} \approx k\mu\delta\left(1-(k-1)\mu\delta\right)$$
$$\approx k\mu\delta$$

pick one of the k calls

one call finishes by the end of our slot.

k-1 calls are still ongoing

These two numbers magically add up to 1. So, we don't have to consider other cases.

To summarize, now we can answer the two questions above via two kinds of probabilities

a) $P[0 \text{ new call}] \approx 1-\lambda\delta$

$P[1 \text{ new call}] \approx \lambda\delta$

b) $P[0 \text{ old-call end}] \approx 1-k\mu\delta$

$P[1 \text{ old-call end}] \approx k\mu\delta$

So, there can be 4 events:

(i) 0 new call & 0 old-call end → K unchanges
(ii) 0 " & 1 " → K − 1
(iii) 1 " & 0 " → K + 1
(iv) 1 " & 1 " → K unchange.

The corresponding probability for each case is

(i) $(1-\lambda\delta)(1-k\mu\delta) \approx 1-\lambda\delta-k\mu\delta$

(ii) $(1-\lambda\delta)(k\mu\delta) \approx k\mu\delta$

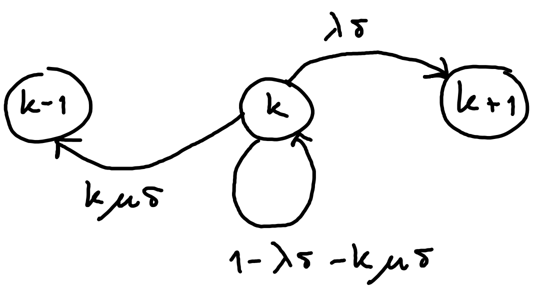
(iii) $(\lambda\delta)(1-k\mu\delta) \approx \lambda\delta$

(iv) $(\lambda\delta)(k\mu\delta) \approx 0$

So, if we have $K=k$ at the beginning of our time slot, then at the end of our time slot, K may

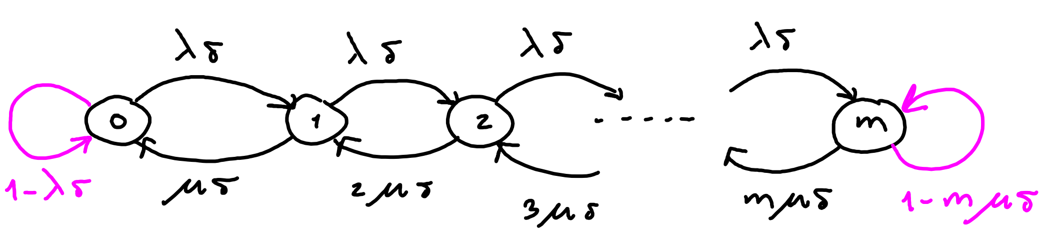
(I) remain unchanged with probability $1-\lambda\delta-k\mu\delta$
(II) decrease by 1 with probability $k\mu\delta$
(III) increase by 1 with probability $\lambda\delta$

This can be summarized in the folling diagram:



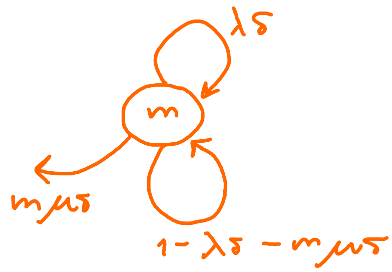
The labels on the arrows indicate transition probabilities (conditional probability of going from one value of K to another value.)

Given m, the possible values of K are 0, 1, 2, ..., m. We can combine the above diagram into one diagram that includes all possible values of K:

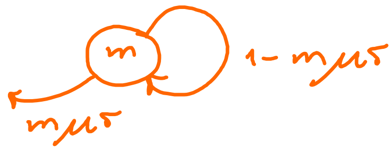


Note that the arrow λδ which should go out of state m will return to state m itself because it means blocked calls which do not increase the value of K.

So we have



By combining the two cyclic arrows, we have



for $k = m$.

The above comple diagram describe a Markov chain!
It is called

Markov chain state diagram for Erlang B

It turns out that, over a long period of time, the system will reach steady state

Rappaport uses Global Balance Equation to get the steady state probabilities.