Ioannis Christofilogiannis 2019030140
Large and Social Networks - 2nd set

Problem 1

The downloads of the 2 files are independent.
Because of exponential distribution, the pdfs will be:

$$f_1(t) = \lambda_1 \cdot e^{-\lambda_1 \cdot t}, \quad f_2(t) = \lambda_2 \cdot e^{-\lambda_2 t}$$

Where t is time until completion

1. $T_1 \sim \exp(\lambda_1)$
   $T_2 \sim \exp(\lambda_2)$
   Expected time until one finishes downloading is: $T = \min(T_1, T_2)$ and its' CDF is:

$$P(T > t) = P(T_1 > t, T_2 > t) = \underbrace{P(T_1 > t) \cdot P(T_2 > t)}_{\text{independence}}$$

$$= e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} = e^{-(\lambda_1+\lambda_2)t}$$

$$\Rightarrow P(T \leq t) = 1 - e^{-(\lambda_1+\lambda_2)t}$$

To get the PDf of T from the CDf:

$$\frac{d}{dt} P(T \leq t) = \frac{d}{dt}\left(1 - e^{-(\lambda_1+\lambda_2)t}\right) = 0 + (\lambda_1+\lambda_2) e^{-(\lambda_1+\lambda_2)t}$$

$$= (\lambda_1+\lambda_2) e^{-(\lambda_1+\lambda_2)t}$$

$$E[T] = \int_0^\infty t f_T(t)\,dt = \int_0^\infty (\lambda_1+\lambda_2)\, t\, e^{-(\lambda_1+\lambda_2)t}\,dt$$

$$= (\lambda_1+\lambda_2) \cdot \frac{1}{(\lambda_1+\lambda_2)^2} = \frac{1}{\lambda_1+\lambda_2}$$ Time until one of the files is downloaded

2. $P(T_1 < T_2) = \int_0^\infty P(T_1 < t \mid t = T_2)\, f_2(t)\, dt$

$$= \int_0^\infty P(t > T_1) \cdot \lambda_2 e^{-\lambda_2 t}\, dt = \int_0^\infty \left(1 - e^{-\lambda_1 t}\right) \cdot \lambda_2 e^{-\lambda_2 t}\, dt$$

$$= \int_0^\infty \left(\lambda_2 e^{-\lambda_2 t} - \lambda_2 e^{-(\lambda_1 + \lambda_2) t}\right) dt$$

$$= \lambda_2 \int_0^\infty \left(e^{-\lambda_2 t} - e^{-(\lambda_1 + \lambda_2) t}\right) dt$$

$$= \lambda_2 \left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1 + \lambda_2}\right) = \lambda_2 \left(\frac{\lambda_1 + \cancel{\lambda_2} - \cancel{\lambda_2}}{\cancel{\lambda_2}(\lambda_1 + \lambda_2)}\right)$$

$= \dfrac{\lambda_1}{\lambda_1 + \lambda_2}$ Probability that file 1 downloads first.

CDF:

$$3.\quad P(t \geq T_1, t \geq T_2) = \underbrace{P(t \geq T_1) \cdot P(t \geq T_2)}_{\text{independence}}$$

$$= \left(1 - e^{-\lambda_1 t}\right)\left(1 - e^{-\lambda_2 t}\right) = 1 - e^{-\lambda_1 t} - e^{-\lambda_2 t} + e^{-(\lambda_1 + \lambda_2)t}$$

PDF:

$$\frac{d}{dt}\left(P(t \geq T_1, t \geq T_2)\right) = \frac{d}{dt}\left(1 - e^{-\lambda_1 t} - e^{-\lambda_2 t} + e^{-(\lambda_1+\lambda_2)t}\right)$$

$$= \lambda_1 e^{-\lambda_1 t} + \lambda_2 e^{-\lambda_2 t} - (\lambda_1 + \lambda_2) e^{-(\lambda_1+\lambda_2)t}$$

$$E[T] = \int_0^\infty t\left(\lambda_1 e^{-\lambda_1 t} + \lambda_2 e^{-\lambda_2 t} - (\lambda_1+\lambda_2) e^{-(\lambda_1+\lambda_2)t}\right) dt$$

$$= \frac{1}{\lambda_1} + \frac{1}{\lambda_2} - \frac{1}{(\lambda_1+\lambda_2)} \quad \text{, like question } a$$

$$= \frac{(\lambda_1+\lambda_2)\lambda_2 + (\lambda_1+\lambda_2)\lambda_1 - \lambda_1\lambda_2}{\lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \lambda_2)}$$

$$= \frac{\lambda_1\lambda_2 + \lambda_2^2 + \lambda_1^2 + \cancel{\lambda_1\lambda_2} - \cancel{\lambda_1\lambda_2}}{\lambda_1\lambda_2(\lambda_1+\lambda_2)} = \frac{\lambda_1^2 + \lambda_2^2 + \lambda_1\lambda_2}{\lambda_1 \cdot \lambda_2 \cdot (\lambda_1 + \lambda_2)}$$

4. Each file has a pdf of: $f(\tau) = \lambda \cdot e^{-\lambda t}$

Then the CDF is:

$$P(T \leq \tau) = P(\underbrace{\max(T_1, T_2, \ldots, T_n)}_{\text{time until everything finishes}} \leq t)$$

$$P(T \leq t) = P(T_1 \leq t) \cdot \underbrace{P(T_2 \leq t) \cdot \ldots \cdot}_{\text{independence}} P(T_n \leq t)$$

We know that $P(T_1 \leq t) = P(T_2 \leq t) = \ldots = P(T_n \leq t)$ $= (1 - e^{-\lambda t})$ so the product is equal to $(1 - e^{-\lambda t})^n$

$$\Rightarrow P(T \leq t) = (1 - e^{-\lambda t})^n$$

And the pdf is: $\frac{d}{dt} P(T \leq t) = \frac{d}{dt} (1 - e^{-\lambda t})^n$

$= n \cdot (1 - e^{-\lambda t})^{n-1} \cdot \lambda \cdot e^{-\lambda t} = n \cdot \lambda \cdot (1 - e^{-\lambda t})^n \cdot e^{-\lambda t}$

(using derivative chain rule)

So with the pdf we can calculate $E[T]$

$$E[T] = \int_0^{+\infty} t \cdot n \cdot \lambda \cdot (1-e^{-\lambda t})^n \cdot e^{-\lambda t}\, dt$$

$$= n \cdot \lambda \cdot \int_0^{+\infty} (1-e^{-\lambda t})^n \cdot e^{-\lambda t}\, dt$$

I did not find a meaningful way to continue this calculation. Using MATLAB or Python could provide us with a good approximation.

Problem 2

1. Delay $\sim Exp(\mu)$ Time between infection and detection

Infection $\sim$ Poisson (λ)

We will calculate the CDF:

$$P(T \leq \underbrace{t-s}) = 1 - e^{-\mu(t-s)}$$

Infection happens on time s and is detected by time t.

2. Infection time s is uniformly distributed over $[0,t]$ because malware infection is a Poisson process, so:

$$p = \frac{1}{t}\int_0^t \left(1 - e^{-\mu(t-s)}\right) ds$$

and: $\int_0^t e^{-\mu(t-s)} ds = \int_t^0 e^{-\mu u}(-du) = \int_0^t e^{-\mu u} du$

$$\left[-\frac{1}{\mu}e^{-\mu u}\right]_0^t = \left[-\frac{1}{\mu}\cdot e^{-\mu t} - \left(-\frac{1}{\mu}e^{0}\right)\right]$$

$$= \frac{1}{\mu}\left(1 - e^{-\mu t}\right)$$

$$\Rightarrow p = \frac{1}{t}\left(\int_0^t ds - \int_0^t e^{-\mu(t-s)}ds\right)$$

$$= \frac{1}{t}\left(t - \frac{1}{\mu}\left(1 - e^{-\mu t}\right)\right)$$

$$= 1 - \frac{1 - e^{-\mu t}}{\mu t}$$

3. $E[N(t)] = \lambda \cdot t$, infections follow Poisson

$N_1(t) = \lambda \cdot t \cdot p$ ← as calculated in (2)

$$\Rightarrow \lambda = \frac{N_1(t)}{t \cdot \left(\frac{1 - e^{-\mu t}}{\mu t}\right)} = \frac{N_1(t) \cdot \mu}{1 - e^{-\mu t}}$$ (using $N_1(t)$ for λ estimate)

4. Estimate of <u>total</u> infections $E[N(t)] = \lambda \cdot t$

so $E[N(t)] = \underbrace{N_1(t)}_{\text{detected}} + \underbrace{N_2(t)}_{\substack{\text{additional} \\ \text{not detected}}}$

So, in total: $N_1(t) + N_2(t) = \lambda \cdot t$

$$\Rightarrow N_1(t) + N_2(t) = \frac{N_1(t) \cdot \mu}{1 - e^{-\mu t}} t$$

$$\Leftrightarrow N_2(t) = N_1(t) \left(\frac{\mu}{1 - e^{-\mu t}} t - 1 \right)$$

Problem 4

1. Using the data from the figure:

$$\begin{cases} \underbrace{R_1}_{\text{input}} = \underbrace{r_1 + 1/3\, R_2}_{\text{output}} & (1) \\ R_2 = r_2 + 1/3\,(R_1 + R_2) & (2) \end{cases}$$

(2) using (1) $\Rightarrow$

$$R_2 = r_2 + 1/3\,(r_1 + 1/3\,R_2 + R_2)$$

$$\Rightarrow R_2 = r_2 + 1/3\,r_1 + 4/9\,R_2$$

$$\overset{r_2=1}{\Longrightarrow} 1/3\,r_1 = 5/9\,R_2 - 1$$

$$\Leftrightarrow \boxed{r_1 = 5/3\,R_2 - 3} \quad (3)$$

$$\Rightarrow R_1 = 5/3\,R_2 - 3 + 1/3\,R_2 = 2R_2 - 3$$

$$\Leftrightarrow \boxed{R_1 = 2R_2 - 3} \quad (4)$$

For stability we must have:

$$\left.\begin{array}{l} R_1 \le \mu_1 \\ R_2 \le \mu_2 \end{array}\right\} \Leftrightarrow \left.\begin{array}{l} 2R_2 - 3 \le 3 \\ R_2 \le 5 \end{array}\right\} \Leftrightarrow \left.\begin{array}{l} R_2 \le 3 \\ R_2 \le 5 \end{array}\right\}$$

So R_2 must be ≤ 3 meaning $R_{2max} = 3$

So using (3) we can calculate r_{1max}:

$$r_{1max} = 5/3\, R_{2max} - 3$$

$$\Leftrightarrow \boxed{r_{1max} = 5 - 3 = 2}$$

2) Setting $r_1 = 0.8\, r_{1max} = 1.6$

from (3): $r_1 = 5/3\, R_2 - 3 \Rightarrow$

$$R_2 = \frac{3}{5}(r_1 + 3) = 2.76$$

From (1): $R_1 = r_1 + 1/3\, R_2 = 1.6 + 1/3 \cdot 2.76 = 2.52$

So given $\rho_i = \frac{\beta_i}{\mu_i}$ we have:

$$\rho_1 = \frac{\beta_1}{\mu_1} = \frac{2.52}{3} = 0.84, \quad \rho_2 = \frac{\beta_2}{\mu_2} = \frac{2.76}{5} = 0.552$$

Both routers are M/M/1, we can use the formula:

$$E[N_1] = \frac{\rho_1}{1-\rho_1} = \frac{0.84}{0.16} = \frac{\cancel{16}\cdot 6}{\cancel{16}} = 6 \qquad (1)$$

$$E[N_2] = \frac{\rho_2}{1-\rho_2} = \frac{0.552}{1-0.552} = 1.23211 \simeq 1.23 \qquad (2)$$

Mean response time for packet entering router 1:

I could not figure out a logical formula in time, that would be a linear combination of (1), (2) like this:

A packet will enter the queue of router 2 with probability 1/3 after router 1

$$E[T_1] = E[N_1]\cdot c_1 + \frac{1}{3}\cdot E[N_2]\cdot c_2$$

c_1, c_2: constants

3. Disregarding the packet's entry point, we can say:

$$\lambda_{total} = r_1 + r_2 = \overset{\text{from (2)}}{1.6} + \overset{\text{known}}{1} = 2.6$$

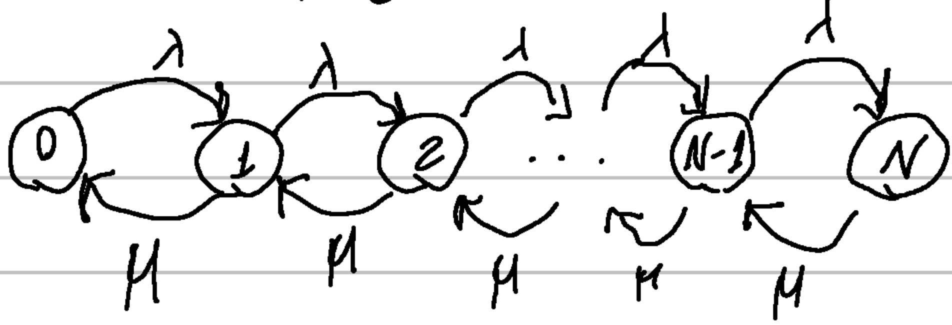
So using the calculations from (2), the mean response time can be calculated as:

$$E[N] = \frac{E[N_1] + E[N_2]}{\lambda_{total}} = \frac{7.23}{2.6} \simeq 2.78$$

This estimation/techique handles the system as a black box.

Proplem 3

1. The CTMC for the described system:



2. Local balance equations:

$$\lambda \pi_0 = \mu \pi_1 \Leftrightarrow \pi_1 = \frac{\lambda}{\mu} \pi_0$$

$$\lambda \pi_1 = \mu \pi_2 \Leftrightarrow \pi_2 = \frac{\lambda}{\mu} \pi_1 = \frac{\lambda}{\mu}\left(\frac{\lambda}{\mu}\pi_0\right) = \left(\frac{\lambda}{\mu}\right)^2 \pi_0$$

So: $\lambda \pi_{N-1} = \mu \pi_N \Leftrightarrow \pi_N = \frac{\lambda}{\mu} \pi_{N-1} = \left(\frac{\lambda}{\mu}\right)^N \pi_0$

$\Rightarrow \pi_N = \left(\frac{\lambda}{\mu}\right)^2 \pi_0$, for $\{0, 1, \dots N\}$

$$\begin{cases} \sum_{n=0}^{N} \pi_n = 1 \\ \lambda / \mu < 1 \text{ (for stability)} \end{cases}$$

$\frac{\lambda}{\mu} = \rho$ $\quad$ Π_0 is a constant $\quad$ from sum

$$\Leftrightarrow \begin{cases} \Pi_0 \cdot \sum_{n=0}^{N} \rho^N = 1 \\ \rho < 1 \end{cases} \Rightarrow \begin{cases} \Pi_0 \cdot \dfrac{1-\rho^{N+1}}{1-\rho} = 1 \\ \rho < 1 \end{cases}$$

$$\Leftrightarrow \Pi_0 = \frac{1-\rho}{1-\rho^{N+1}}, \rho < 1$$

$$\Rightarrow \Pi_n = \rho^n \cdot \Pi_0 = \rho^n \cdot \frac{1-\rho}{1-\rho^{N+1}}$$

3. Utilization = not in state 0 (inactive)

Util = $1 - \Pi_0$ (aka probability of existing in any other state)

$$\text{Util} = 1 - \frac{1-\rho}{1-\rho^{N+1}} = \frac{\cancel{1} - \rho^{N+1} - \cancel{1} + \rho}{1-\rho^{N+1}}$$

$$= \frac{\rho - \rho^{N+1}}{1-\rho^{N+1}} \quad \text{and} \quad \rho = \frac{\lambda}{\mu}$$

4. Loss probability is equivalent to the chain being in state N, so π_N.

$$P_{loss} = \pi_N = \rho^N \cdot \frac{1-\rho}{1-\rho^{N+1}}, \rho = \frac{\lambda}{\mu}$$

from (2).

5. The rate is equal to the rate of jobs arriving (independent to the system and equal to λ) while the system is in state N (π_N from (4)).

$$\text{rate\_loss} = \lambda \cdot P_{loss} = \lambda \cdot \rho^N \cdot \frac{1-\rho}{1-\rho^{N+1}}$$

$$= \frac{\lambda^{N+1}}{\mu^N} \cdot \frac{1-\left(\frac{\lambda}{\mu}\right)}{1-\left(\frac{\lambda}{\mu}\right)^{N+1}}$$ if we substitute ρ again.

$$6.\ E[N] = \sum_{n=0}^{N} n \cdot \pi_N = \underset{\text{constant}}{\pi_0} \cdot \sum_{n=0}^{N} n \cdot \rho^n$$

$$= \pi_0 \cdot \rho \cdot \sum_{n=0}^{N} n \cdot \rho^{n-1} = \pi_0 \cdot \rho \cdot \sum_{n=0}^{\infty} \frac{d}{d\rho} (\rho^n)$$

$$= \pi_0 \cdot \rho \ \frac{d}{d\rho} \sum_{n=0}^{\infty} \rho^n$$

$$= \pi_0 \cdot \rho \cdot \frac{d}{d\rho} \left(\frac{1-\rho^{N+1}}{1-\rho} \right)$$

order of sum, derivative can be changed and $\frac{d}{dp} p^n = n p^{n-1}$

$$= \pi_0 \cdot \rho \ \frac{(N+1)\rho^N (\rho-1) - (\rho^{N+1}-1) \cdot 1}{(1-\rho)^2}$$

$$= \frac{(N \cdot \rho^N + \rho^N)(\rho - 1) - \rho^{N+1} + 1}{(1-\rho)^2} \cdot \pi_0 \cdot \rho$$

$$= \frac{N\rho^{N+1} - \cancel{\rho^{N+1}} + \cancel{\rho^{N+1}} - N\rho^N - \rho^N}{(1-\rho)^2} \cdot \pi_0 \cdot \rho$$

$$= \frac{N\rho^{N+1} - (N+1)\rho^{N} + 1}{(1-\rho)^2} \cdot \pi_0 \cdot \rho$$

$$\Rightarrow E[N] = \frac{\cancel{1-\rho}}{1-\rho^{N+1}}\rho \cdot \frac{N\rho^{N+1} - (N+1)\rho^{N} + 1}{(1-\rho)^{\cancel{2}}}$$

$$= \frac{N\rho^{N+2} - (N+1)\rho^{N+1} + \rho}{(1-\rho^{N+1})(1-\rho)}$$

7. $E[T]$ for only the jobs that enter the system. We will use little's law

$$E[T] = \frac{E[N]}{\lambda_{eff}} = \frac{E[N]}{\lambda(1-P_{block})}$$

$$\Rightarrow E[T] = \frac{N\rho^{N+2} - (N+1)\rho^{N+1} + \rho}{(1-\rho^{N+1})(1-\rho)\cdot\lambda\left(1 - \frac{\rho^{N}(1-\rho)}{1-\rho^{N+1}}\right)}$$

$$\Rightarrow E[T] = \frac{N\rho^{N+2} - (N+1)\rho^{N+1} + \rho}{\cancel{(1-\rho^{N+1})}(1-\rho)\left(\frac{1-\rho^{N+1} - \rho^N(1-\rho)}{\cancel{1-\rho^{N+1}}}\right)}$$

$$\Rightarrow E[T] = \frac{N\rho^{N+2} - (N+1)\rho^{N+1} + \rho}{(1-\rho)(1-\rho^N)}$$

8. $N = 4$, $\rho = \lambda/\mu = 0.8$

$P_{loss} = \rho^N \cdot \frac{1-\rho}{1-\rho^{N+1}} \simeq 0.122$ (calculator)

P_{loss} (double buffer) $= (0.8)^8 \cdot \frac{0.2}{1-(0.8)^9} \simeq 0.0388$
$(N_2 = 2N = 8)$

P_{loss} (double CPU speed) $= (0.4)^4 \cdot \frac{0.6}{1-(0.4)^5} \simeq 0.0155$
$(\mu_2 = 2\mu \Rightarrow \rho_2 = \rho/2 = 0.4)$
service time is halved

We see that doubling the CPU speed is more effective because Ploss (2·CPU) is less than half of Ploss (2·buffer)

9. $N = 4$, $\rho = 0.4$:

We don't need to compute original Ploss

Ploss (double buffer) $= (0.4)^8 \cdot \dfrac{0.6}{1-(0.4)^9} \simeq 0.000393$
$(N_2 = 2N = 8)$

Ploss (double CPU speed) $= (0.2)^4 \cdot \dfrac{0.8}{1-(0.2)^5}$
$(\mu_2 = 2\mu \Rightarrow \rho_2 = \rho/2 = 0.2)$
service time is halved

$\simeq 0.00128$

In this problem doubling the buffer is the right choise, as Ploss is magnitudes smaller.

10. In the first scenario, utilization is very high $\rho = 0.8$, so lowering the utilization by half by doubling CPU speed is more effective.

In the second scenario utilization is already low $\rho = 0.4$, so doubling the buffer we get a larger queue, lowering the loss/quit probability more drastically, while doubling the CPU would make the utilization even smaller which would not be as beneficial.

We know that the delay increases drastically when util is large, like 0.8 so these results make sense.

In general we need a CPU upgrade when the exponential growth of delay starts to show, with large utilization and buffer in other cases.