A Mathematical Tutorial on Poisson Counter Driven Stochastic Differential Equations

For study group presentation by Adrian

June 4 and 28, 2005
updated on June 28, 2005

1 Overview of Differential Calculus

- A *continuous* function $f(x)$ has independent variable x
  - Differentiation of $f(x)$ with respect to x is denoted by $\frac{df(x)}{dx}$
  - Differential of $f(x)$ is an approximation of the change in $f(x)$: $df(x) \approx f(x+\delta x) - f(x)$
  - Differentials can be related by the derivative: $df(x) = \frac{df(x)}{dx}dx$
- A continuous multi-variable function $f(x_1, x_2, \dots, x_n)$ has independent variables $x_1, x_2, \dots, x_n$
  - Differentiation of $f(\cdot)$ with respect to x_k is denoted by $\frac{\partial f}{\partial x_k}$
  - Differential of $f(\cdot)$ is expressed as a sum: $df = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} dx_k$
  - Think: The term $\frac{\partial f}{\partial x_k} dx_k$ is the change in f due to dx_k amount of change in a single variable x_k
- Inverse of differentiation is integration

$$
\begin{aligned}
& dx = A(x,t)dt \\
\Longrightarrow \quad & x(t) = x(0) + \int_0^t A(x,\tau) d\tau
\end{aligned}
$$

2 Introduction to SDE

2.1 SDE with respect to a Wiener process [4, 7]

- A standard Wiener process/Brownian motion W_t
  - Infinitesimal increments dW_t in time dt has density: $\frac{1}{\sqrt{2\pi dt}} e^{-dW_t^2/2dt}$
  - Mean of dW_t: $\overline{dW_t} = E\{dW_t\} = 0$
  - Variance of $dW_t = E\left\{(dW_t - \overline{dW_t})^2\right\} = E\left\{(dW_t)^2\right\} = \overline{(dW_t)^2}$
    - But according to the density function, the variance is dt, hence

$$\overline{(dW_t)^2} = dt \tag{1}$$

  - A function of Wiener process: $f(t, W_t)$ has the differential

$$df(t, W_t) = f(t+dt, W_t + dW_t) - f(t, W_t) \tag{2}$$

- Taylor's expansion: [5]

$$f(x_1,\ldots,x_n)=\sum_{j=0}^{\infty}\left\{\frac{1}{j!}\left[\sum_{k=1}^{n}(x_k-a_k)\frac{\partial}{\partial x'_k}\right]^j(x'_1,\ldots,x'_n)\right\}_{x'_1=a_1,\ldots,x'_n=a_n}$$

$$\begin{aligned}f(x+\delta x,y+\delta y)&=f(x,y)+\left[\frac{\partial f(x,y)}{\partial x}\delta x+\frac{\partial f(x,y)}{\partial y}\delta y\right]\\&\quad+\frac{1}{2!}\left[\frac{\partial^2 f(x,y)}{\partial x^2}(\delta x)^2+2\frac{\partial^2 f(x,y)}{\partial x\partial y}\delta x\delta y+\frac{\partial^2 f(x,y)}{\partial y^2}(\delta y)^2\right]\\&\quad+\cdots+\frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k}\frac{\partial^n f(x,y)}{\partial x^k\partial y^{n-k}}(\delta x)^k(\delta y)^{n-k}+\cdots\end{aligned}$$

  - Compare: Taylor's expansion for one-variable function: $f(x)=\sum_{k=0}^{\infty}\frac{(x-a)^k}{k!}f^{(k)}(a)$
  - Hence tyhe Taylor's expansion of (2):

$$\begin{aligned}df(t,W_t)&=-f(t,W_t)+f(t+dt,W_t+dW_t)\\&=-f(t,W_t)+f(t,W_t)+\frac{\partial f(t,W_t)}{\partial t}dt+\frac{\partial f(t,W_t)}{\delta W_t}dW_t\\&\quad+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial t^2}(dt)^2+\frac{\partial^2 f(t,W_t)}{\partial t\partial W_t}dtdW_t+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2}(dW_t)^2+\cdots\\&=\frac{\partial f(t,W_t)}{\partial t}dt+\frac{\partial f(t,W_t)}{\delta W_t}dW_t+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial t^2}(dt)^2+\frac{\partial^2 f(t,W_t)}{\partial t\partial W_t}dtdW_t+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2}(dW_t)^2+\cdots\end{aligned}$$

    * Substituting (1), the mean behavior of $df(t,W_t)$ is therefore:

$$\begin{aligned}df(t,W_t)&=\frac{\partial f(t,W_t)}{\partial t}dt+\frac{\partial f(t,W_t)}{\delta W_t}dW_t+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial t^2}(dt)^2+\frac{\partial^2 f(t,W_t)}{\partial t\partial W_t}dtdW_t+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2}(dW_t)^2+\cdots\\&=\frac{\partial f(t,W_t)}{\partial t}dt+\frac{\partial f(t,W_t)}{\delta W_t}dW_t+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial t^2}(dt)^2+\frac{\partial^2 f(t,W_t)}{\partial t\partial W_t}dtdW_t+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2}dt+\cdots\\&\approx\frac{\partial f(t,W_t)}{\partial t}dt+\frac{\partial f(t,W_t)}{\delta W_t}dW_t+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2}dt\\&=\left[\frac{\partial f(t,W_t)}{\partial t}+\frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2}\right]dt+\frac{\partial f(t,W_t)}{\delta W_t}dW_t\qquad\text{as } dt\to 0\end{aligned}$$

- Chain rule:
  - Let $dx(t,W_t)=a(x,t)dt+b(x,t)dW_t$
  - For $f(x(t,W_t))$,

$$\begin{aligned}df(x(t,W_t))&=\left[a(x,t)\frac{df(x)}{dx}+\frac{1}{2}b^2(x,t)\frac{d^2f}{dx^2}\right]dt+b(x,t)\frac{df}{dx}dW_t\\\because\quad dx(t,W_t)&=\left[\frac{\partial x(t,W_t)}{\partial t}+\frac{1}{2}\frac{\partial^2 x(t,W_t)}{\partial W_t^2}\right]dt+\frac{\partial x(t,W_t)}{\partial W_t}dW_t\\\therefore\quad df(x(t,W_t))&=\left[\left(\frac{\partial x(t,W_t)}{\partial t}+\frac{1}{2}\frac{\partial^2 x(t,W_t)}{\partial W_t^2}\right)\frac{df(x)}{dx}+\frac{1}{2}\left(\frac{\partial x(t,W_t)}{\delta W_t}\right)^2\frac{d^2f(x)}{dx^2}\right]dt+\frac{\partial x(t,W_t)}{\partial W_t}\frac{df(x)}{dx}dW_t\end{aligned}$$

  - This is called the one-dimensional Itô's formula, named after Kiyoshi Itô (伊藤清)

- Traditional product rule:

$$\begin{aligned}d(f(t)g(t))&=f(t+dt)g(t+dt)-f(t)g(t)\\&=[f(t+dt)-f(t)]\,g(t+dt)+f(t)\,[g(t+dt)-g(t)]\\&=[f(t+dt)-f(t)]\,[g(t+dt)-g(t)]+[f(t+dt)-f(t)]\,g(t)+f(t)\,[g(t+dt)-g(t)]\\&=df(t)dg(t)+g(t)df(t)+f(t)dg(t)\\&=\frac{df(t)}{dt}\frac{dg(t)}{dt}(dt)^2+\frac{df(t)}{dt}g(t)dt+\frac{dg(t)}{dt}f(t)dt\end{aligned}$$

  - If dt is infinitesimal, $(dt)^2 = 0$ and we get

$$d(f(t)g(t)) = \frac{df(t)}{dt}g(t)dt + \frac{dg(t)}{dt}f(t)dt$$

- Product rule in It

$$\begin{aligned}
d(f(t,W_t)g(t,W_t)) &= f(t+dt,W_t+dW_t)g(t+dt,W_t+dW_t) - f(t,W_t)g(t,W_t) \\
&= [f(t+dt,W_t+dW_t) - f(t,W_t)]\,g(t+dt,W_t+dW_t) + f(t,W_t)\,[g(t+dt,W_t+dW_t) - g(t,W_t)] \\
&= [f(t+dt,W_t+dW_t) - f(t,W_t)]\,[g(t+dt,W_t+dW_t) - g(t,W_t)] \\
&\quad + [f(t+dt,W_t+dW_t) - f(t,W_t)]\,g(t,W_t) + f(t,W_t)\,[g(t+dt,W_t+dW_t) - g(t,W_t)] \\
&= df(t,W_t)dg(t,W_t) + g(t,W_t)df(t,W_t) + f(t,W_t)dg(t,W_t) \\
\because \quad df(t,W_t) &= \left(\frac{\partial f(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2}\right)dt + \frac{\partial f(t,W_t)}{\partial W_t}dW_t = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}\right)dt + \frac{\partial f}{\partial W_t}dW_t \\
dg(t,W_t) &= \left(\frac{\partial g(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 g(t,W_t)}{\partial W_t^2}\right)dt + \frac{\partial g(t,W_t)}{\partial W_t}dW_t = \left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial W_t^2}\right)dt + \frac{\partial g}{\partial W_t}dW_t \\
\therefore \quad d(f(t,W_t)g(t,W_t)) &= \left[\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}\right)dt + \frac{\partial f}{\partial W_t}dW_t\right]\left[\left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial W_t^2}\right)dt + \frac{\partial g}{\partial W_t}dW_t\right] \\
&\quad + g\left[\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}\right)dt + \frac{\partial f}{\partial W_t}dW_t\right] + f\left[\left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial W_t^2}\right)dt + \frac{\partial g}{\partial W_t}dW_t\right] \\
&= \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}\right)\left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial W_t^2}\right)(dt)^2 + \frac{\partial f}{\partial W_t}\left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial W_t^2}\right)dtdW_t \\
&\quad + \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}\right)\frac{\partial g}{\partial W_t}dtdW_t + \frac{\partial f}{\partial W_t}\frac{\partial g}{\partial W_t}(dW_t)^2 \\
&\quad + g\left[\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}\right)dt + \frac{\partial f}{\partial W_t}dW_t\right] + f\left[\left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial W_t^2}\right)dt + \frac{\partial g}{\partial W_t}dW_t\right] \\
(\text{as } t \to 0) &= \frac{\partial f}{\partial W_t}\frac{\partial g}{\partial W_t}(dW_t)^2 + g\left[\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}\right)dt + \frac{\partial f}{\partial W_t}dW_t\right] + f\left[\left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial W_t^2}\right)dt + \frac{\partial g}{\partial W_t}dW_t\right] \\
&= \frac{\partial f}{\partial W_t}\frac{\partial g}{\partial W_t}dt + \left[\left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial W_t^2}\right)g + \left(\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial W_t^2}\right)f\right]dt + \left[\frac{\partial f}{\partial W_t}g + \frac{\partial g}{\partial W_t}f\right]dW_t
\end{aligned}$$

which yields the following Itô's product rule:

$$\begin{aligned}
&d\left(f(t,W_t)g(t,W_t)\right) = \\
&\left[\left(\frac{\partial f(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 f(t,W_t)}{\partial W_t^2}\right)g(t,W_t) + \left(\frac{\partial g(t,W_t)}{\partial t} + \frac{1}{2}\frac{\partial^2 g(t,W_t)}{\partial W_t^2}\right)f(t,W_t) + \frac{\partial f(t,W_t)}{\partial W_t}\frac{\partial g(t,W_t)}{\partial W_t}\right]dt \\
&\qquad + \left[\frac{\partial f(t,W_t)}{\partial W_t}g(t,W_t) + \frac{\partial g(t,W_t)}{\partial W_t}f(t,W_t)\right]dW_t
\end{aligned}$$

- Rationale:

$$dx = a(x,t)dt + b(x,t)dW_t$$
$$x(t) = x(0) + \int_0^t a(x,t)dt + \int_0^t b(x,t)dW_t$$

The first equation is the stochastic differential equation. Itô's solution is to find a function $x(t)$ that satisfy the second equation. Which we call it the solution of the first one.

2.2 SDE with respect to a Poisson counting process [1]

- Poisson Counter Driven Stochastic differential equations is of the form:

$$dX(t) = f(X(t))dt + \sum_{i=1}^{n} g_i(X(t))dN_i(t)$$

where $\{X(t)\}$ is a stochastic process described by the stochastic differential equation, $N_i(t)$ are Poisson counters that drives $X(t)$, and $f(x)$, $g_i(x)$ are real-valued functions

- If it is driven by Wiener process, $N_i(t)$ becomes $W_i(t)$
- Poisson process is in discrete domain, i.e. $N_i(t) \in \mathbb{N}$; but Wiener process is in continuous domain $W_i(t) \in \mathbb{R}$. This causes the PCDSDE differ from the previous section

- Poisson counter $N_i(t)$ with $dN_i(t) = 0$ if no event i occur at t and $dN_i(t) = 1$ if event $i$occur at t, i.e.

$$dN_i = \begin{cases} 1 & \text{at Poisson arrival} \\ 0 & \text{elsewhere} \end{cases} \tag{3}$$
$$E[dN_i] = \lambda_i dt$$

- Properties of stochastic differential equations

1. cádlág: continue á droite, limite á gauche. If $N(t)$ jumps at τ

$$\lim_{t\to\tau^-} X(t) = X(\tau^-)$$
$$X(\tau) = X(\tau^-) + g(X(\tau^-))$$

2. If $h(t) = h(X(t))$ is a function of a stochastic process, then due to the jump nature:

$$\begin{aligned} dh(t) &= \frac{dh}{dX}\left(f(X(t))\,dt + \sum_{i=1}^{n} g_i(X(t))\,dN_i(t)\right) \\ &= \frac{dh}{dX} f(X(t))\,dt + \sum_{i=1}^{n} \frac{dh}{dX} g_i(X(t))\,dN_i(t) \\ &= \frac{dh}{dX} f(X(t))\,dt + \sum_{i=1}^{n} \left[h(X(t) + g_i(X(t))) - h(X(t))\right] dN_i(t) \end{aligned} \tag{4}$$

3. Let λ_i be the rate associated with $N_i(t)$, then

$$dE[X(t)] = E[f(X(t))]dt + \sum_{i=1}^{n} E[g_i(X(t))]\lambda_i dt$$
$$\text{or} \quad \frac{dE[X(t)]}{dt} = E[f(X(t))] + \sum_{i=1}^{n} \lambda_i E[g_i(X(t))] \tag{5}$$

3 Examples

3.1 M/G/1 Queue [1,2]

- Imagine a M/G/1 queue with server capacity of 1
- Arrival is represented by Poisson counting process $\{N(t)\}$ with arrival rate λ, general service time per customer is X
- Let $W(t)$ be the amount of work in the system (which can also be the queueing time of the customer arriving at t), then

$$\begin{aligned} dW(t) &= \begin{cases} -dt + X\,dN(t) & W(t) > 0 \\ X\,dN(t) & W(t) = 0 \end{cases} \\ &= -\mathbf{1}(W(t) > 0)dt + X\,dN(t) \end{aligned}$$

- By (5), we have

$$\begin{aligned} \frac{dE[W(t)]}{dt} &= -E[\mathbf{1}(W(t) > 0)] + \lambda E[X] \\ &= -\Pr[W(t) > 0] + \lambda E[X] \end{aligned}$$

- If the system is ergodic and stable, $\rho \triangleq \lambda E[X] < 1$ and $dE[W(t)]/dt = 0$, hence

$$\begin{aligned} \frac{dE[W(t)]}{dt} &= 0 \\ \therefore \quad -\Pr[W(t) > 0] + \lambda E[X] &= 0 \\ \Pr[W(t) > 0] &= \lambda E[X] = \rho \end{aligned}$$

- Similarly, the differential for the second moment of $W(t)$:

$$\begin{aligned}
dW^2(t) &= \frac{dW^2}{dW}\left(-\mathbf{1}(W(t)>0)\right)dt + \left[(W(t)+X)^2 - W^2(t)\right]dN(t) \\
&= -2W(t)\mathbf{1}(W(t)>0)dt + \left[W^2(t)+2XW(t)+X^2-W^2(t)\right]dN(t) \\
&= -2W(t)\mathbf{1}(W(t)>0)dt + \left[2XW(t)+X^2\right]dN(t) \\
\frac{dE[W^2(t)]}{dt} &= -2E[W(t)\mathbf{1}(W(t)>0)] + \lambda\left(2E[W(t)X]+E[X^2]\right) \\
&= -2E[W(t)] + \lambda\left(2E[W(t)]E[X]+E[X^2]\right) \\
&= -2E[W(t)] + 2\rho E[W(t)] + \lambda E[X^2] \\
&= 2E[W(t)](\rho-1) + \lambda E[X^2]
\end{aligned}$$

- In steady state, $dE[W^2(t)]/dt = 0$ which yields the Pollaczek-Khinchin formula

$$\begin{aligned}
0 &= 2E[W(t)](\rho-1) + \lambda E[X^2] \\
2(1-\rho)E[W(t)] &= \lambda E[X^2] \\
E[W(t)] = E[W_q] &= \frac{\lambda E[X^2]}{2(1-\rho)}
\end{aligned}$$

3.2 Window size in TCP and determining the rate [6]

- The paper describes the TCP traffic model by Poisson counters N_{TD} (triple duplicate ACKs) and N_{TO} (timeout)
- Window size W is described by the differential equation:

$$dW = \frac{dt}{RTT} - \frac{W}{2}dN_{TD} + (1-W)dN_{TO} \tag{6}$$

- Taking expectation on (6) will have:

$$\begin{aligned}
E[dW] &= \frac{dt}{RTT} - \frac{E[W]}{2}E[dN_{TD}] + (1-E[W])E[dN_{TO}] \\
dE[W] &= \frac{dt}{RTT} - \frac{E[W]}{2}\lambda_{TD}dt + (1-E[W])\lambda_{TO}dt \\
\frac{dE[W]}{dt} &= \frac{1}{RTT} - \frac{E[W]}{2}\lambda_{TD} + (1-E[W])\lambda_{TO} \\
&= \frac{1}{RTT} + \lambda_{TO} - \left(\lambda_{TO} + \frac{\lambda_{TD}}{2}\right)E[W]
\end{aligned}$$

since

$$\begin{aligned}
& y'(t) = a + by(t) \\
\Longrightarrow \qquad & y(t) = -\frac{a}{b} + Ce^{bt}
\end{aligned}$$

therefore

$$\begin{aligned}
E[W] &= \frac{\frac{1}{RTT}+\lambda_{TO}}{\lambda_{TO}+\frac{\lambda_{TD}}{2}} + C\exp\left(-\left(\lambda_{TO}+\frac{\lambda_{TD}}{2}\right)t\right) && \exists C \in \mathbb{R} \\
&= \frac{\frac{1}{RTT}+\lambda_{TO}}{\lambda_{TO}+\frac{\lambda_{TD}}{2}} && \text{as } t \to \infty \text{ (steady state)}
\end{aligned}$$

- The throughtput R is the expected window size divided by the RTT, hence

$$R = \frac{1}{RTT}\left(\frac{\frac{1}{RTT}+\lambda_{TO}}{\lambda_{TO}+\frac{\lambda_{TD}}{2}}\right)$$

which we can determine the data rate of TCP by giving RTT, timeout rate and triple-ACK rate

- In [6], a further refinement of the window size W is done

- Maximum window size allowed is $W_{\max} = K$. Hence no more additive increase after $W = K$
- Equation (6) is therefore rewritten as:

$$dW = I_M(W)\frac{dt}{RTT} - \frac{W}{2}dN_{TD} + (1-W)dN_{TO}$$

where

$$I_M(W) = \begin{cases} 1 & W < M \\ 0 & W = M \end{cases}$$

- Similaring doing expectation,

$$\begin{aligned}
E[dW] &= E[I_M(W)]\frac{dt}{RTT} - \frac{E[W]}{2}E[dN_{TD}] + (1-E[W])E[dN_{TO}] \\
dE[W] &= E[I_M(W)]\frac{dt}{RTT} - \frac{E[W]}{2}\lambda_{TD}dt + (1-E[W])\lambda_{TO}dt \\
\frac{dE[W]}{dt} &= \Pr[W<M]\frac{1}{RTT} - \frac{E[W]}{2}\lambda_{TD} + (1-E[W])\lambda_{TO} \\
0 &= \Pr[W<M]\frac{1}{RTT} + \lambda_{TO} - \left(\lambda_{TO} + \frac{\lambda_{TD}}{2}\right)E[W] \\
E[W] &= \frac{\Pr[W<M]\frac{1}{RTT} + \lambda_{TO}}{\lambda_{TO} + \frac{1}{2}\lambda_{TD}} \\
&= \frac{(1-\Pr[W=M])\frac{1}{RTT} + \lambda_{TO}}{\lambda_{TO} + \frac{1}{2}\lambda_{TD}}
\end{aligned}$$

- Please read the paper for the rest, the final result is:

$$P[W=M] = \frac{2\lambda_{TO}^2 + 2\lambda_{TO} + \lambda_{TO}\lambda_{TD} + 2\lambda_{TO}\frac{1}{RTT} + 2\frac{1}{RTT^2} + 2\frac{1}{RTT}}{(\frac{1}{RTT}+1)(2M\lambda_{TO} + M\lambda_{TD} + 2\frac{1}{RTT})}$$

3.3 Adrian's Unpublished Research Work

- A pipe of capacity 1 with two types of traffic: TCP and UDP
- Each TCP user arrives in Poisson process with rate λ_t and UDP user arrives with rate λ_u
- TCP user has file size in exponential distribution with mean $1/\mu_t$; UDP user requires a fixed bandwidth $\alpha < 1$ but it stay in the network for an exponential time with mean $1/\mu_u$
- UDP user will refuse to enter the network if the network is full, i.e. if the network has n TCP and m UDP, and the following satisfied

$$n\varepsilon + (m+1)\alpha > 1$$

then the UDP user will not enter the network

- Admission function is defined as: $I(n,m) = \begin{cases} 0 & n\varepsilon + (m+1)\alpha > 1 \\ 1 & n\varepsilon + (m+1)\alpha \le 1 \end{cases}$
- Let $W(t)$ be the total number of bytes to be transferred by the network at time t
- PCDSDE:

$$dW(t) = -\mathbf{1}(W>0)dt + S_t dN_t + I(n,m)S_u dN_u$$

  - S_t and S_u are random variables describing the data size (number of bytes) of a newly arriving TCP and UDP user, respectively
  - N_t and N_u are Poisson counters decribing the arrival of TCP and UDP users
  - The expected value of $I(n,m)$ denotes the ratio of admitted UDP users to the total UDP arrivals

- Taking the mean:

$$\begin{aligned}
dW(t) &= -\mathbf{1}(W>0)dt + S_t dN_t + I(n,m)S_u dN_u \\
dE[W] &= -E[\mathbf{1}(W>0)]dt + E[S_t dN_t] + E[I(n,m)S_u dN_u] \\
&= -\Pr[W>0]dt + E[S_t]E[dN_t] + E[I(n,m)]E[S_u]E[dN_u] \\
&= -\Pr[W>0]dt + \frac{1}{\mu_t}\lambda_t dt + \Pr[n\varepsilon+(m+1)\alpha>1]\frac{\alpha}{\mu_u}\lambda_u dt \\
\frac{dE[W]}{dt} &= -\Pr[W>0] + \frac{\lambda_t}{\mu_t} + \Pr[n\varepsilon+(m+1)\alpha>1]\alpha\frac{\lambda_u}{\mu_u}
\end{aligned}$$

But if the system is in steady state, the mean workload $E[W]$ should be independent of the time t, hence $dE[W]/dt = 0$ is expected, so we get

$$\begin{aligned}
0 &= -\Pr[W>0] + \frac{\lambda_t}{\mu_t} + \Pr[n\varepsilon+(m+1)\alpha>1]\alpha\frac{\lambda_u}{\mu_u} \\
\Pr[n\varepsilon+(m+1)\alpha>1]\alpha\frac{\lambda_u}{\mu_u} &= \Pr[W>0] - \frac{\lambda_t}{\mu_t} \\
\Pr[n\varepsilon+(m+1)\alpha>1] &= \frac{\Pr[W>0]-\rho_t}{\alpha\rho_u} \qquad (7)
\end{aligned}$$

with $\rho_t = \lambda_t/\mu_t$ and $\rho_u = \lambda_u/\mu_u$. The term $\Pr[W>0]$ is the percentage of busy time of the pipe, which can be approximated by $\Pr[W>0] \approx \min(1, \alpha\rho_u + \rho_t)$.

- The second moment of $W(t)$:

$$\begin{aligned}
\because \quad dW(t) &= -\mathbf{1}(W>0)dt + S_t dN_t + I(n,m)S_u dN_u \\
\therefore \quad dW^2(t) &= \frac{dW^2}{dW}\left(-\mathbf{1}(W>0)\right)dt + \left[(W(t)+S_t)^2 - W^2(t)\right]dN_t + \left[(W(t)+I(n,m)S_u)^2 - W^2(t)\right]dN_u \qquad (8) \\
&= -2W(t)\mathbf{1}(W>0)dt + \left[W^2(t)+2S_tW(t)+S_t^2-W^2(t)\right]dN_t + \left[W^2(t)+2I(n,m)S_uW(t)+I^2(n,m)S_u^2-W^2(t)\right]dN_u \\
&= -2W(t)\mathbf{1}(W>0)dt + \left[2S_tW(t)+S_t^2\right]dN_t + + \left[2I(n,m)S_uW(t)+I(n,m)S_u^2\right]dN_u \\
\frac{dE[W^2]}{dt} &= -2E[W] + \lambda_t\left(2E[S_tW]+E[S_t^2]\right) + E[I(n,m)]\lambda_u\left(2E[S_uW]+E[S_u^2]\right) \\
&= -2E[W] + \lambda_t\left(2E[S_t]E[W]+E[S_t^2]\right) + E[I(n,m)]\lambda_u\left(2E[S_u]E[W]+E[S_u^2]\right) \\
&= -2E[W] + \lambda_t\left(2\frac{1}{\mu_t}E[W]+E[S_t^2]\right) + E[I(n,m)]\lambda_u\left(2\frac{\alpha}{\mu_u}E[W]+E[S_u^2]\right) \qquad (9)
\end{aligned}$$

  - Second moment of exponential variable:

$$\begin{aligned}
E[X^2] &= \int_0^\infty X^2\lambda e^{-\lambda X}dX \\
&= -\int_0^\infty X^2 de^{-\lambda X} \\
&= -\left[X^2e^{-\lambda X}\right]_0^\infty + \int_0^\infty 2Xe^{-\lambda X}dX \\
&= \frac{-2}{\lambda}\int_0^\infty X de^{-\lambda X} \\
&= \frac{-2}{\lambda}\left[Xe^{-\lambda X}\right]_0^\infty + \frac{2}{\lambda}\int_0^\infty e^{-\lambda X}dX \\
&= \frac{2}{\lambda}\left[\frac{-1}{\lambda}e^{-\lambda X}\right]_0^\infty \\
&= \frac{2}{\lambda^2}
\end{aligned}$$

  - Hence substitute into (9) and setting the derivative to be zero, we have

$$\begin{aligned}
0 &= -2E[W] + \lambda_t\left(2\frac{1}{\mu_t}E[W]+\frac{2}{\mu_t^2}\right) + E[I(n,m)]\lambda_u\left(2\frac{\alpha}{\mu_u}E[W]+\frac{2\alpha^2}{\mu_u^2}\right) \\
E[W] &= \rho_tE[W] + \frac{\rho_t}{\mu_t} + E[I(n,m)]\alpha\rho_uE[W] + \frac{R\alpha^2\rho_u}{\mu_u}
\end{aligned}$$

$$(1-\rho_t - E[I(n,m)]\alpha\rho_u)E[W] = \frac{\rho_t}{\mu_t} + \frac{E[I(n,m)]\alpha^2\rho_u}{\mu_u}$$
$$E[W] = \frac{\frac{\rho_t}{\mu_t} + \frac{E[I(n,m)]\alpha^2\rho_u}{\mu_u}}{1-\rho_t - E[I(n,m)]\alpha\rho_u}$$

after substituting $E[I(n,m)]$ as in (7), it becomes

$$\begin{aligned}
E[W] &= \frac{\frac{\rho_t}{\mu_t} + \frac{E[I(n,m)]\alpha^2\rho_u}{\mu_u}}{1-\rho_t - E[I(n,m)]\alpha\rho_u} \\
&= \left(\frac{\rho_t}{\mu_t} + \frac{\alpha^2\rho_u}{\mu_u}\frac{\Pr[W>0]-\rho_t}{\alpha\rho_u}\right) \Big/ \left(1-\rho_t - \frac{\Pr[W>0]-\rho_t}{\alpha\rho_u}\alpha\rho_u\right) \\
&= \left(\frac{\rho_t}{\mu_t} + \frac{\alpha(\Pr[W>0]-\rho_t)}{\mu_u}\right) \Big/ (1-\rho_t - \Pr[W>0]+\rho_t) \\
&= \frac{\rho_t\mu_u + \alpha\Pr[W>0] - \alpha\lambda_t}{\mu_t\mu_u} \Big/ (1-\Pr[W>0]) \\
&= \frac{\rho_t\mu_u + \alpha\Pr[W>0] - \alpha\lambda_t}{\mu_t\mu_u(1-\Pr[W>0])}
\end{aligned}$$

- Because the pipe has capacity 1, using Little's formula, the mean population of TCP traffic is:

$$\begin{aligned}
n &= \lambda \times T \\
&= \lambda_t \frac{\rho_t\mu_u + \alpha\Pr[W>0] - \alpha\lambda_t}{\mu_t\mu_u(1-\Pr[W>0])} \\
&= \frac{\rho_t\mu_u + \alpha\Pr[W>0] - \alpha\lambda_t}{\rho_t\mu_u(1-\Pr[W>0])} \\
&\approx \frac{\rho_t\mu_u + \alpha\rho' - \alpha\lambda_t}{\rho_t\mu_u(1-\rho')}
\end{aligned}$$

- So, what's next?!

3.4 Fluid Queue [3]

- A fluid queue with serving capacity c and a buffer size B feed by markov on-off source
- When the source is in "on" state, fluid arrives with rate $h > c$; no arrival in "off" state
- There is a threshold K such that, when the fluid in queue exceeds K while the source just turns to "on" state, the whole arrival is discarded
  - Rationale: An IP packet (usually >1 kbyte) is divided into multiple 53-byte cells in ATM. One single ATM cell drop will cause the whole IP packet unrecoverable. If we model IP packet arrival as markov on-off source and the fluid queue as ATM link, this behavior is optimal to use the bandwidth
- The source is in silence period for an exponentially distributed time with mean $1/\lambda_1$; the burst size is also exponentially distributed with mean $1/\lambda_2$
- Let $\theta \in \{0,1\}$ denotes the behavior of the source and $x \in \{0,1\}$ denotes the arrival pattern with discarding policy, i.e. if $x = \theta = 1$, the fluid queue is accepting fluid arrival with rate h; if $x = \theta = 0$, the queue is clearing and no arrial to the queue; if $x = 0$ while $\theta = 1$, the arrival is discarded due to the buffer level exceeded K *at the time* θ turns from 0 to 1.
- The process of buffer level $v(t)$ and arrival pattern $x(t)$ are denoted by

$$\begin{aligned}
dx(t) &= [1-x(t)]\,\mathbf{1}(v(t)<K)dN_1 - x(t)dN_2 \\
dv(t) &= -c\mathbf{1}(v(t)>0)dt + hx(t)\mathbf{1}(v(t)<B)dt + cx(t)\mathbf{1}(v(t)=B)dt
\end{aligned}$$

  where N_1 is a Poisson counter with rate λ_1 (signals the end of silence period) and N_2 is a Poisson counter with rate λ_2 (signals the end of bursty arrival)
  - The first term of $dx(t)$ means arrival of a burst is accepted only if $v(t) < K$ and $x(t) = 0$
    The second term means the termination of a burst will cause $x(t)$ to change only if $x(t) = 1$

- The first term of $dv(t)$ means that if the queue is not empty, the buffer is cleared at the rate of c
  The second term means if $x(t) = 1$, the queue will be filled with the rate of h unless $v(t) \geq B$, which buffer overflow will occur
  The third term means, if buffer overflow occur, and $x(t) = 1$, $v(t)$ will keep at B

- The differential of the $(n+1)$-th power of $v(t)$, using equation (4):

$$\begin{aligned} dv^{n+1} &= (n+1)v^n dv \\ &= (n+1)v^n \left[-c\mathbf{1}(v > 0)dt + hx\mathbf{1}(v < B)dt + cx\mathbf{1}(v = B)dt\right] \\ dv^n x &= nxv^{n-1} dv + v^n dx \\ &= nxv^{n-1} \left[-c\mathbf{1}(v > 0)dt + hx\mathbf{1}(v < B)dt + cx\mathbf{1}(v = B)dt\right] \\ &\quad + v^n \left[(1-x)\mathbf{1}(v < K)dN_1 - xdN_2\right] \end{aligned}$$

and taking the expectation:

$$\begin{aligned} dE[v^{n+1}] &= (n+1)\left(-cE[v^n\mathbf{1}(v > 0)]dt + hE[v^n x\mathbf{1}(v < B)]dt + cE[v^n x\mathbf{1}(v = B)]dt\right) \\ &= (n+1)\left(-cE[v^n]dt + hE[v^n x|v < B]\Pr[v < B]dt + cE[v^n x|v = B](1 - \Pr[v < B])dt\right) \\ &= (n+1)\left(hE[v^n x|v < B]\Pr[v < B]dt - cE[v^n x|v < B]\Pr[v < B]dt\right) \\ dE[v^n x] &= n\left(-cE[v^{n-1}x\mathbf{1}(v > 0)] + hE[v^{n-1}x^2\mathbf{1}(v < B)] + cE[v^{n-1}x^2\mathbf{1}(v = B)]\right)dt \\ &\quad + E[v^n(1-x)\mathbf{1}(v < K)]\lambda_1 dt - E[v^n x]\lambda_2 dt \\ &= n\left(-cE[v^{n-1}x] + hE[v^{n-1}x|v < B]\Pr[v < B] + cE[v^{n-1}x|v = B](1 - \Pr[v = B])\right)dt \\ &\quad + E[v^n(1-x)\mathbf{1}(v < K)]\lambda_1 dt - E[v^n x]\lambda_2 dt \end{aligned}$$

- Notes:
  1. $x^2 = x$ because $x \in \{0, 1\}$
  2. $E[v^n x^m \mathbf{1}(v > 0)] = E[v^n x^m]$
  3. $v(t)$ is upperbounded by B. Hence $\Pr[v < B] = 1 - \Pr[v = B] \stackrel{\Delta}{=} 1 - p_2$
  4. $E[v^n x^m] = E[v^n x^m|v < B]p_2 + E[v^n x^m|v = B](1 - p_2)$
  5. We define $\Pr[v < K] \stackrel{\Delta}{=} p_1 < p_2$
  6. If $v = B$, x must be 1. Hence $E[v^n x^m|v = B] = E[v^n|v = B] = B^n$

- Therefore we can rewrite the differential $dE[v^n x]$ into:

$$\begin{aligned} dE[v^n x] &= n\left(-cE[v^{n-1}x] + hE[v^{n-1}x|v < B]p_2 + cE[v^{n-1}x|v = B](1 - p_2)\right)dt \\ &\quad + E[v^n(1-x)|v < K]\lambda_1 p_1 dt - E[v^n x]\lambda_2 dt \\ &= n\left(-cE[v^{n-1}x] + h\left(E[v^{n-1}x] - E[v^{n-1}x|v = B](1 - p_2)\right) + cE[v^{n-1}x|v = B](1 - p_2)\right)dt \\ &\quad + E[v^n(1-x)|v < K]\lambda_1 p_1 dt - E[v^n x]\lambda_2 dt \\ &= n\left((h-c)E[v^{n-1}x] + (h-c)E[v^{n-1}x|v = B](1 - p_2)\right)dt \\ &\quad + E[v^n(1-x)|v < K]\lambda_1 p_1 dt - E[v^n x]\lambda_2 dt \\ &= n\left((h-c)E[v^{n-1}x|v < B]p_2\right)dt \\ &\quad + E[v^n - v^n x)|v < K]\lambda_1 p_1 dt - E[v^n x]\lambda_2 dt \\ &= n\left((h-c)E[v^{n-1}x|v < B]p_2\right)dt \\ &\quad + E[v^n|v < K]\lambda_1 p_1 dt - E[v^n x|v < K]\lambda_1 p_1 dt - E[v^n x|v < B]p_2\lambda_2 dt - E[v^n x|v = B](1 - p_2)\lambda_2 dt \\ &= n\left((h-c)E[v^{n-1}x|v < B]p_2\right)dt \\ &\quad + E[v^n|v < K]\lambda_1 p_1 dt - E[v^n x|v < K]\lambda_1 p_1 dt - E[v^n x|v < B]p_2\lambda_2 dt - B^n(1 - p_2)\lambda_2 dt \end{aligned}$$

- Substitute the following:

$$\begin{aligned} E[v^n|v < B]p_2 &= E[v^n|v < K]p_1 + E[v^n|K \leq v < B](p_2 - p_1) \\ E[v^n x|v < B]p_2 &= E[v^n x|v < K]p_1 + E[v^n x|K \leq v < B](p_2 - p_1) \end{aligned}$$

into the differentials $dE[v^n x]$ and $dE[v^{n+1}]$ will get:

$$dE[v^{n+1}] = (n+1)\left(hE[v^n x|v < B]\Pr[v < B]dt - cE[v^n x|v < B]\Pr[v < B]dt\right)$$

$$
\begin{aligned}
&= (n+1)\left(hE[v^n x|v<B]p_2 dt - cE[v^n x|v<B]p_2 dt\right) \\
&= (n+1)\left(hE[v^n x|v<K]p_1 + hE[v^n x|K \le v<B](p_2-p_1) - cE[v^n|v<K]p_1 - cE[v^n|K \le v<B](p_2-p_1)\right)dt \\
dE[v^n x] &= n\left((h-c)E[v^{n-1}x|v<B]p_2\right)dt \\
&\quad + E[v^n|v<K]\lambda_1 p_1 dt - E[v^n x|v<K]\lambda_1 p_1 dt - E[v^n x|v<B]p_2\lambda_2 dt - B^n(1-p_2)\lambda_2 dt \\
&= n\left((h-c)E[v^{n-1}x|v<K]p_1 + (h-c)E[v^{n-1}x|K \le v<B](p_2-p_1)\right)dt \\
&\quad + E[v^n|v<K]\lambda_1 p_1 dt - E[v^n x|v<K]\lambda_1 p_1 dt \\
&\quad - E[v^n x|v<K]p_1\lambda_2 dt - E[v^n x|K \le v<B](p_2-p_1)\lambda_2 dt - B^n(1-p_2)\lambda_2 dt.
\end{aligned}
$$

- So setting the two differentials to zero (as the steady behavior, the expectation is independent of time t), we get:

$$
\begin{aligned}
0 &= hE[v^n x|v<K]p_1 + hE[v^n x|K \le v<B](p_2-p_1) - cE[v^n|v<K]p_1 - cE[v^n|K \le v<B](p_2-p_1) \\
0 &= n(h-c)p_1 E[v^{n-1}x|v<K] + n(h-c)(p_2-p_1)E[v^{n-1}x|K \le v<B] \\
&\quad + E[v^n|v<K]\lambda_1 p_1 - E[v^n x|v<K]\lambda_1 p_1 \\
&\quad - E[v^n x|v<K]p_1\lambda_2 - E[v^n x|K \le v<B](p_2-p_1)\lambda_2 - B^n(1-p_2)\lambda_2 \\
&= n(h-c)E[v^{n-1}x] - n(h-c)E[v^{n-1}x|v=B](1-p_2) - B^n(1-p_2)\lambda_2 \\
&\quad - E[v^n x|v<K]p_1(\lambda_1+\lambda_2) - E[v^n x|K \le v<B](p_2-p_1)\lambda_2 + E[v^n|v<K]\lambda_1 p_1
\end{aligned}
$$

or in matrix form:

$$
\begin{pmatrix} hp_1 & h(p_2-p_1) & -cp_1 & -c(p_2-p_1) \\ -p_1(\lambda_1+\lambda_2) & -\lambda_2(p_2-p_1) & \lambda_1 p_1 & 0 \end{pmatrix}
\begin{pmatrix} E[v^n x|v<K] \\ E[v^n x|K \le v<B] \\ E[v^n|v<K] \\ E[v^n|K \le v<B] \end{pmatrix}
+ \begin{pmatrix} 0 \\ n(h-c)E[v^{n-1}x] - n(h-c)B^{n-1} - B^n(1-p_2)\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
$$

- Eliminating $E[v^n x|v<K]$ from the above yields:

$$
\begin{aligned}
\therefore E[v^n|v<K] &= \\
&\Bigg[\frac{h}{c(\lambda_1+\lambda_2)}\left(n(h-c)E[v^{n-1}x] - B^n(1-p_2)\left(n(h-c)B^{-1}+\lambda_2\right) + (p_2-p_1)\lambda_1 E[v^n x|K \le v<B]\right) \\
&\qquad - (p_2-p_1)E[v^n|K \le v<B]\Bigg]\left(p_1 - \frac{\lambda_1 p_1 h}{c(\lambda_1+\lambda_2)}\right)^{-1}
\end{aligned}
$$

- In [3], it can finally derive a formula for $E[v^n]$ and then by using

$$
\mathscr{L}\{\Pr[v]\} = E[e^{-sv}] = E[\sum_{n=0}^{\infty}\frac{(-sv)^n}{n!}] = \sum_{n=0}^{\infty}\frac{(-s)^n}{n!}E[v^n]
$$

to derive the Laplace transform of the buffer size distribution, which can be used to find the probability distribution and other useful things.

References

[1] Roger W. Brockett. Stochastic control. Lecture Notes. Harvard University, Cambridge, MA., 1983. Provided by the author.

[2] Roger W. Brockett, Weibo Gong, and Yang Guo. Stochastic analysis for fluid queueing systems. In *Proceedings of the 38th IEEE Conference on Decision and Control*, volume 3, pages 3077–3082, December 1999.

[3] Parijat Dube and Eitan Altman. On the workload process in a fluid queue with a responsive bursty input and selective discarding. *Journal of Probability in the Engineering and Informational Sciences*, 17:527–543, 2003.

[4] Lawrence Craig Evans. An introduction to stochastic differential equations. Lecture Notes. Department of Mathematics, U.C. Berkeley. version 1.2, http://math.berkeley.edu/ evans/.

[5] MathWorld. Taylor series. http://mathworld.wolfram.com/TaylorSeries.html.

[6] Vishal Misra, Weibo Gong, and Don Towsley. Stochastic differential equation modeling and analysis of tcp windowsize behavior. Technical Report ECE-TR-CCS-99-10-01, Department of Electrical and Computer Engineering, University of Massachusetts, October 1999. Presented at Performance 99, Istanbul.

[7] Sheldon M. Ross. *Introduction to Probability Models*. Academic Press, 8th edition, 2003.