Closed Form Solution

March 18, 2014

For simplicity, let's assume that we are working with a net with only 1 attacker "A" and 1 node that can possibly be infected "B". This makes it simple because there are only 2 possible "infection orderings". One is when the node gets infected before the observation window ends at T and one where it doesn't get infected. The latter is trivial to compute. First some notation, Let λ_1 be the rate that "B" sends messages given it is not infected. Let λ_2 be the rate that "B" sends messages given it is infected. Let γ be the rate at which "A" sends infected messages to "B". Let k_x be the number of messages that "B" sends before x and let N be the total number of messages that "B" sends. Since only one node can be infected drop the subscript and let $\bar{z}$ be the infection time of "B". Here we go...

$$\int_0^T P(\text{ data } |\bar{z})P(\bar{z}|s)d\bar{z}$$

$$= \int_0^T P(\text{ messages before } \bar{z}|\bar{z}) \times P(\text{ messages after } \bar{z}|\bar{z}) \times P(\bar{z}|s)d\bar{z} \tag{1}$$

$$= \int_0^T \frac{(\lambda_1\bar{z})^{k_{\bar{z}}}e^{-\lambda_1\bar{z}}}{k_{\bar{z}}!} \times \frac{(\lambda_2(T-\bar{z}))^{N-k_{\bar{z}}}e^{-\lambda_2(T-\bar{z})}}{(N-k_{\bar{z}})!} \times \gamma e^{-\gamma\bar{z}}d\bar{z} \tag{2}$$

Denote m_i as the time of the ith message sent by "B" $\tag{3}$

and $m_0 = 0$ and $m_{N+1} = T$ $\tag{4}$

$$= \sum_{i=0}^N \int_{m_i}^{m_{i+1}-\Delta t} \frac{(\lambda_1\bar{z})^{k_{\bar{z}}}e^{-\lambda_1\bar{z}}}{k_{\bar{z}}!} \times \frac{(\lambda_2(T-\bar{z}))^{N-k_{\bar{z}}}e^{-\lambda_2(T-\bar{z})}}{(N-k_{\bar{z}})!} \times \gamma e^{-\gamma\bar{z}}d\bar{z} \tag{5}$$

Note that for each integral, k is a constant. This is because we subtract an
arbitrarily small positive number Δt, from the upper limit of the integral

Therefore, let k_i be the number of messages sent by node "B" at or before m_i

$$= \lim_{\Delta t \to 0^+} \sum_{i=0}^N \frac{1}{k_i!(N-k_i)!} \int_{m_i}^{m_{i+1}-\Delta t} (\lambda_1\bar{z})^{k_I}e^{-\lambda_1\bar{z}} \times (\lambda_2(T-\bar{z}))^{N-k_i}e^{-\lambda_2(T-\bar{z})} \times \gamma e^{-\gamma\bar{z}}d\bar{z}$$

$$= \lim_{\Delta t \to 0^+} \sum_{i=0}^N \frac{\lambda_1^{k_i}\lambda_2^{N-k_i}}{k_i!(N-k_i)!} \int_{m_i}^{m_{i+1}-\Delta t} \bar{z}^{k_I}e^{-\lambda_1\bar{z}} \times (T-\bar{z})^{N-k_i}e^{-\lambda_2(T-\bar{z})} \times \gamma e^{-\gamma\bar{z}}d\bar{z}$$

$$= \lim_{\Delta t \to 0^+} \sum_{i=0}^N \frac{e^{-\lambda_2 T}\lambda_1^{k_i}\lambda_2^{N-k_i}}{k_i!(N-k_i)!} \int_{m_i}^{m_{i+1}-\Delta t} \bar{z}^{k_I}e^{-\lambda_1\bar{z}} \times (T-\bar{z})^{N-k_i}e^{\lambda_2\bar{z}} \times \gamma e^{-\gamma\bar{z}}d\bar{z}$$

$$= \lim_{\Delta t \to 0^+} \sum_{i=0}^N \frac{e^{-\lambda_2 T}\lambda_1^{k_i}\lambda_2^{N-k_i}\gamma}{k_i!(N-k_i)!} \int_{m_i}^{m_{i+1}-\Delta t} \bar{z}^{k_I}(T-\bar{z})^{N-k_i}e^{\bar{z}(\lambda_2-\lambda_1-\gamma)}d\bar{z}$$

The issue with this integral is that the rates are non-homogeneous. If the rates were
homogeneous, the integration would be able to be done through integration by parts.

However, the $(T-\bar{z})^{N-k_i}$ term makes this integral "difficult"

However, when I do not know how to solve an integral, I turn to something that can, namely, Mathematica

You can see the exact Mathematica command below but for any two intervals points m, m_2

The final value is

$$= \lim_{\Delta t \to 0^+} \sum_{i=0}^N \frac{e^{-\lambda_2 T}\lambda_1^{k_i}\lambda_2^{N-k_i}\gamma}{k_i!(N-k_i)!} e^{(-\lambda_2+\lambda_1+\gamma)}T^{N+2}\left[\text{Beta}(\frac{m_i}{T}, 2+k_i, 1-k_i+n) - \text{Beta}(\frac{m_{i+1}}{T}, 2+k_i, 1-k_i+n)\right]$$

$$\tag{6}$$

where Beta is the incomplete Beta function with parameters (x, a, b)

The main problem with this is that I have **no idea** how to get to that evaluation of the integral. So going forward I see 3 paths. The first is to attack this closed form calculation another way. Maybe I approached it the wrong way. The second path is to actually try and work toward that integral and try and derive it by hand. The third path is to go ahead and implement it computationally and see what it looks like.

As a note, another reason why this integral is so difficult is because the bounds on the integral are the times that the possibly infected node sends messages. Therefore, it is possible that the node gets infected between times that it sends messages. This makes me think that if instead of having m_i as the messages sent by node "B", we define m_i as the time of ith message defined by the union of all messages sent by "B" and all possible infected messages sent to "B". In that case, we wouldn't need to do a definite integral over possible infection times because we know that a node can only become infected at the time it receives a malicious message. This might be the possible fourth path.

Mathematica command :

```
{Integrate[zbar^ki Exp (zbar*(lambda2 - lambda1 - gamma)) (T - zbar)^(n - ki), {zbar, m1, m2},

Assumptions -> {T > 0, n > 0, n >=  ki, ki >= 0,  gamma > 0, lambda2 > lambda1 > 0, m1 > 0,

m2 > 0, m2 > m1, m2 <  T, m1 <= T}]}
```