Stochastic Formulation of Scalability and Quality-of-Information Satisfiability in Wireless Networks

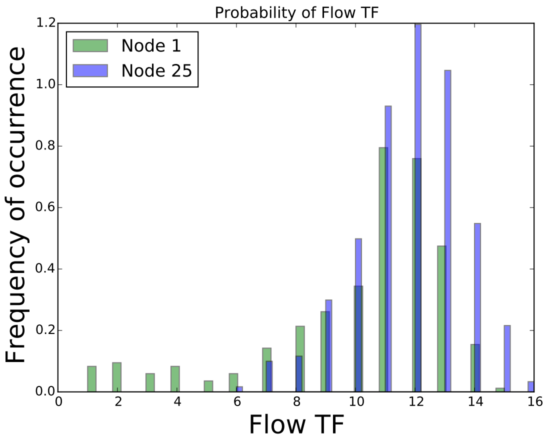


Fig. 1. Simulation results: Distribution of max TF of flows originating in the first node and the middle node in a line network with 50 nodes. Image size = 360 KB

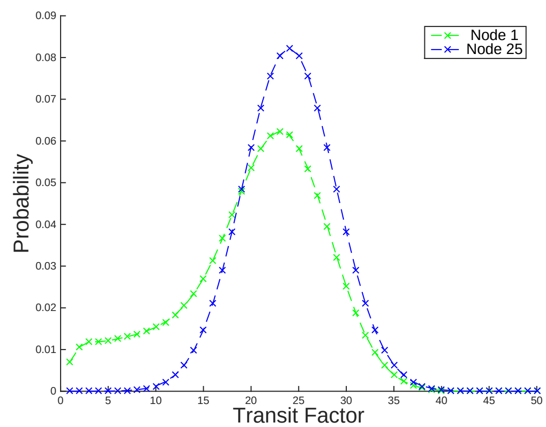
I. Traffic Factor

I ran some simulations to see empirically what the actual maximum traffic factor is for flows. Here, each node keeps track of the number of flows that it is currently forwarding, and every packet in each flow captures the maximum TF of any node along the path that it travels. The result for flows originating in the first node in the line and the middle node in the line are plotted in 1.

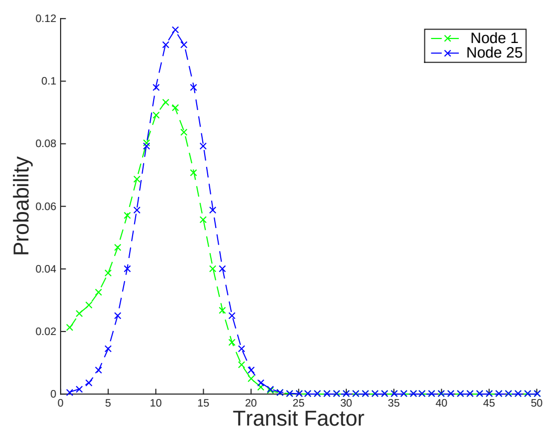
In my working document of deriving the analytical model, I came up with the following expression for the expected TF in a node x:

$$\frac{2 \cdot (x-1) \cdot (N-x)}{N-1} \tag{1}$$

conditioning this expression on all possible destinations and the corresponding node with the max TF, we get the distributions for the TF of flows of the first and middle nodes plotted in Figure 2(a). Clearly, the mean value is much higher than what I was seeing in the simulations. Interestingly enough, it seems that the mean values were about twice as high as I was seeing in simulations for several different network sizes, so I computed the distribution using the above expression divided by 2. That plot is in Figure 2(b), which actually resembles the empirical results pretty well. That leads me to my first question: Did I make a mistake in deriving an expression for TF that would have made it off by a factor of 2??



(a) Analysis results: Distribution of expected maximum TF values for flows in the first and middle nodes using the derived TF expression in the other document



(b) Analysis results: Distribution of expected maximum TF values for flows in the first and middle nodes dividing the derived TF expression in the other document by 2

Fig. 2. We can observe the empirical statistical properties of the Traffic Factor for each node's position in a line network (here with 125 nodes).

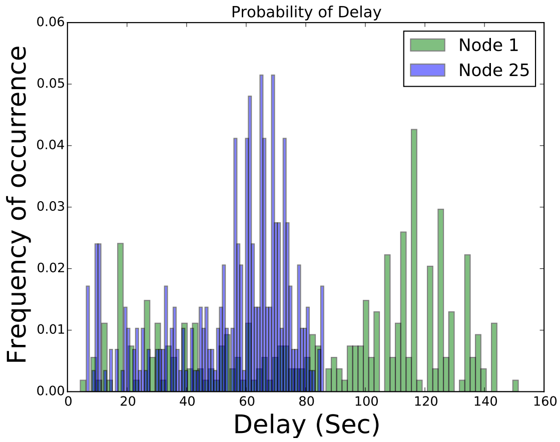


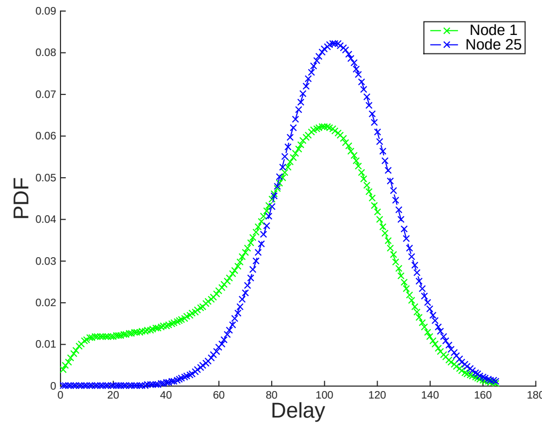
Fig. 3. Simulation results: Distribution of delays of flows originating in the first node and the middle node in a line network with 50 nodes. Image size = 360 KB

II. Delay

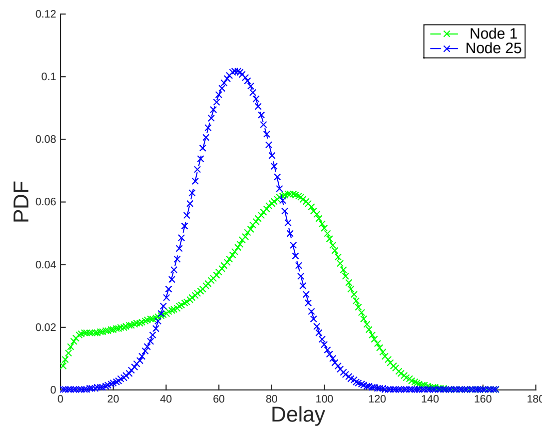
I also collected and plotted a histogram of the delays that flows from each of the nodes experience in the simulations. Figure 3 shows this distribution of delays for the same nodes again. Here, we can see that the flows that experience the long delays are those that start in the end node. This is not exactly the argument made in the WCNC paper, which accurately deduces that the maximum TF occurs at the center node, but also uses an average path length of $N/4$, which is for flows originating from the center node. I'm actually still trying to reconcile why all of the results that use that formula were so accurate, but it seems that if we wrong on that equation, then it probably wasn't too far off because it does seem to be at least close.

Coming back to the formulation of a delay distribution, though, I also computed and plotted analytical values of the derived model. First, if I use the derived expressions from the other document, including the expression for TF that seems to have a mean twice as large as the empirical values, the delays don't seem correct. They are not always close to ending around the maximum delays observed in simulations, and they don't project the flows from node 1 having higher delay than those of the middle node. To provide an example, I plotted this case in Figure 2(a).

So, I first tried using the distribution of TFs that is closer to the simulation results (in Figure 2(b)), but that doesn't provide matching distributions, because the flows from node 1 do not exceed the center node as seen in the simulations. It seems that path lenght has a larger impact on delays than what I had previously formulated (more evidence of this in Figure **??**)...so I started playing with the factors. For a few scenarios, multiplying the second portion of the delay, i.e., the multi-hop propagation portion, by a factor proportional to the number of packets in a flow seems to be close(r). Figure 2(b) calculates the delay distribution using the original TF expression divided 2 and multiplying the $C_2 \cdot PL(i,j)$ part of the delay by $\frac{P_n}{2}$,



(a) Analysis results: Distribution of delay values for flows in the first and middle nodes using the derived TF expression in the other document



(b) Analysis results: Distribution of delay values for flows in the first and middle nodes dividing the derived TF expression in the other document by 2

Fig. 4. We can observe the empirical statistical properties of the Traffic Factor for each node's position in a line network (here with 125 nodes).

where P_n is the number of packets in a flow.

In the WCNC paper, I argue that this delay is only experienced once per flow, and, thus, should not be multiplied by this extra factor, but I'm wondering if that also isn't quite accurate?? I'm still trying to run more simulations and trace more closely what is happening to figure out where the discrepancies are and what is causing them.

III. Other Simulation Results

Some other things that I tracked in simulations and plotted provide some insight, too. First, I modified the original stochastic expression of delay because I realized that the Traffic Factor and path length are not really independent. If you think about a flow beginning in node 1, this makes sense since the max TF of the flow depends exactly on what the destination is. Figure 5 shows a scatter plot of the max TF experienced by each flow given its path length. Flows beginning in the center node are pretty evenly distributed and

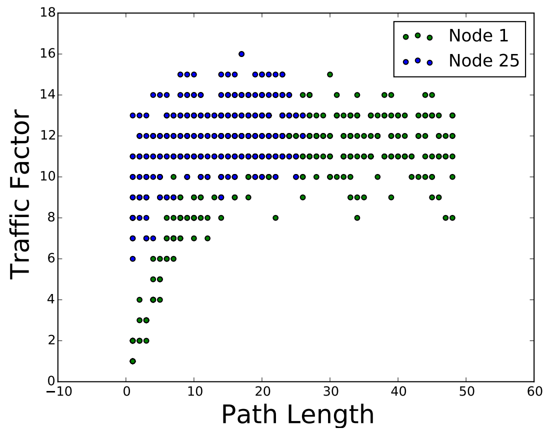


Fig. 5. Simulation results: Scatter of TF vs. Path Length for flows in the first node and the middle node in a line network with 50 nodes. Image size = 360 KB

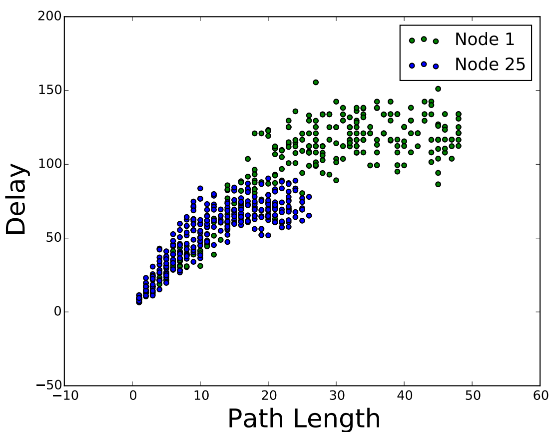


Fig. 6. Simulation results: Scatter of delay vs. Path Length for flows in the first node and the middle node in a line network with 50 nodes. Image size = 360 KB

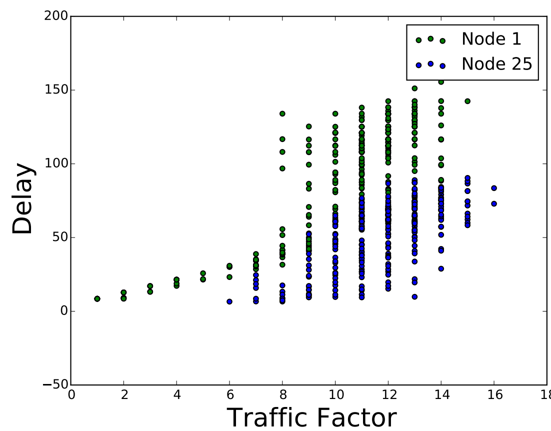


Fig. 7. Simulation results: Scatter of delay vs. max Traffic Factor for flows in the first node and the middle node in a line network with 50 nodes. Image size = 360 KB

appear to be somewhat independent of path length. The flows beginning in node 1, though, have a linear dependence on path length until it reaches about half of the network size, at which point it levels off and is evenly distributed independently for path lengths to the full length of the network.

As I stated earlier, it seems that the path length has a larger impact on overall delay than I had accounted for previously. This seems to be reinforced by plotting the delay vs. path length of flows in Figure 6. The trend is pretty clearly linear. I also plotted the delay vs. traffic factor values in Figure 7. This relationship also seems to be linear, but more spread. Since I believe that path length and traffic factor are not independent, I'm not exactly sure how to interpret these results, but I think it might point to a larger scaling factor for the multi-hop propagation component of delay than we originally surmised.