Modeling TCP Congestion with Tropical Algebra

Waylon Jepsen,

Department of Computer Science, College of Natural Sciences, Colorado State University

Max Plus Algebra

Introduction

Prior work showed that TCP could be modeled by Max plus Tropical Algebra with some assumptions. In this paper, we assess the impact of those assumptions on our ability to use Max-Plus algebra to analyze actual TCP traffic. We validate the model against simulated network traffic to fine granularity.

The TCP Max-Plus Model Max-Plus algebra is the semi-ring structure over real numbers with operations $\oplus$ an $\otimes$ where the $\oplus$ operator performs the max operation such that $x \oplus y = max(x, y)$, and the $\otimes$ operator is the traditional addition operation such that $x \otimes y = x + y$ where $x, y \in \mathbb{R} \cup \{-\infty\}$. The model applies to a single source sending packets to a single destination over a path of K routers in series. Each router is assumed to have a FIFO queue. The n^{th} packet arriving at router i experiences an *aggregated service time* denoted $\sigma_i(n)$ [1]. This model defines a propagation delay between routers where the delay for router i to router j denoted $d_{i,j}$ is the time taken by a packet to travel from i to router j.

Network State Let $x_i(n)$ be the time at which packet n starts its service time at router i. Then define $y_i(n) = x_i(n) + \sigma_i(n)$ as the time when packet n leaves router i. Let v_n be the window size experienced by packet $n-1$. The window size experienced by packet n, v_n is generally not the same as the actual window size w_n. If we know the sequence $\{v_n\}$ then the set $\{y_i(n)\}$ satisfies the following equations for a sequence of routers K.

$$y_0(n) = y_K(n - v_{n-1} \otimes d_{K,0}) \tag{1}$$

$$y_i(n) = [y_{i-1}(n) \otimes d_{i-1,i} \oplus y_i(n-1)] \otimes \sigma_i(n), i = \{1, ..., K\} \tag{2}$$

Let $Y(n) = (y_1(n), y_2(n), ..., y_K(n) \in \mathbb{R}^{1,K}_{max}$ be the sequence of times that packet n leaves all routers $i, .., K$. Now let $Z(n) = (Y(n), Y(n-1), ..., Y(n-w^*+1))^t \in \mathbb{R}^{1,K}_{max}$. The models shows that if a system is initially empty, and if the sequence of experienced window sizes $\{V_n\}$ is known, then the dater vectors satisfy the following equation [1].

$$Z(n) = A_{v_{n-1}}(n) \otimes Z(n-1), n \geq 1, Z(0) = (0, ..., 0)^t \tag{3}$$

$Z(n)$ can be thought of as a sequence of packet traces $\{Y(n), Y(n-1), ..., Y(n-w^*+1)\}$ and w^* is the maximum window size. The $A_{v_{n-1}}$ terms are defined in terms of M and M' as $A_{w^*}(n) = (M(n)|\epsilon|...|\epsilon|M'(n)) \oplus D$. Where M and M' are defined as

$$(M(n))_{ij} = \begin{cases} \sum_{k=j}^{i} \sigma_k(n) \oplus \sum_{k=j}^{i-1} d_{k,k+1}, & i \geq j \\ -\infty, & i > j \end{cases}$$

$$(M'(n))_{ij} = \begin{cases} \sum_{k=1}^{i} (d_{k,k-1} \oplus \sigma_k(n)) \oplus d_{K,0}, & j = K \\ -\infty, & j < K \end{cases}$$

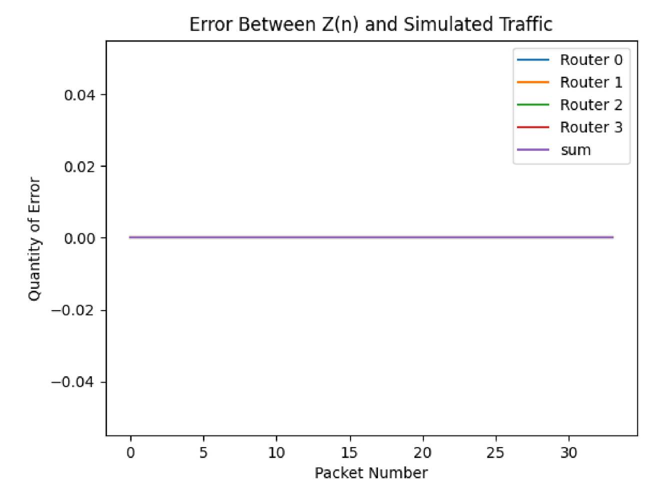
and D is the square matrix of dimension Kw^* with all its entries equal to $-\infty$ except $D_{K+i,i}, i = 1, .., K(w^*-1)$ which are equal to zero.

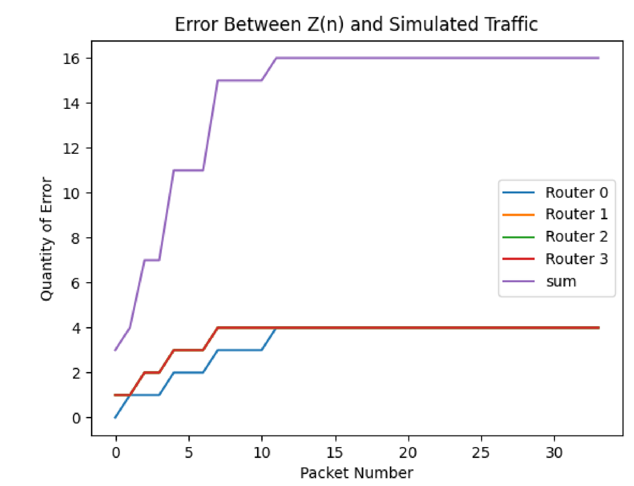
Simulation and Validation

The simulations maintained network granularity accounting for the link and switch bandwidth and packet size. The model was configured for a linear topology and T-bone topology with multiple TCP connections. For each TCP connection, the simulated data was validated against both equations §2 and §3. Packet size variance for TCP connections α and β was tested: the control $\alpha = \beta$, random sampling from exponential distribution, $\alpha = 2\beta$, and alternating α between $\frac{1}{4}\beta$ and $\frac{1}{3}\beta$.

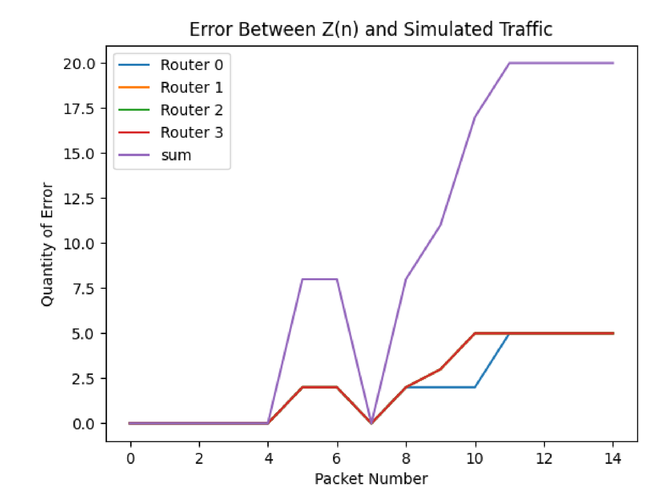
Results

Control We can see that when $\alpha = \beta$, the α connection wins a race condition and has no error, while the other is displaced with an error proportional to the size of α.

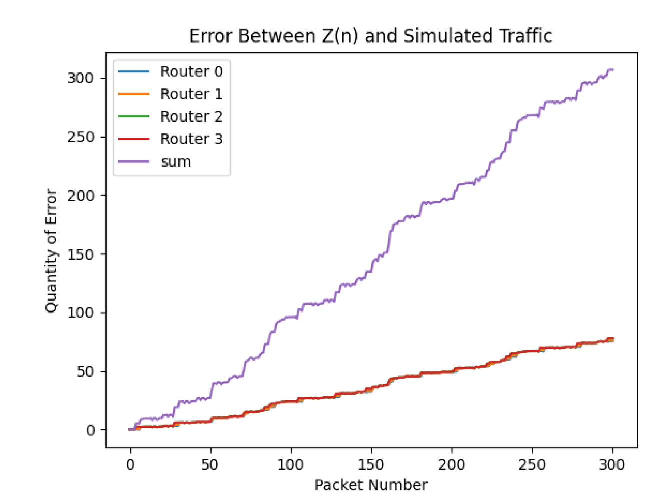


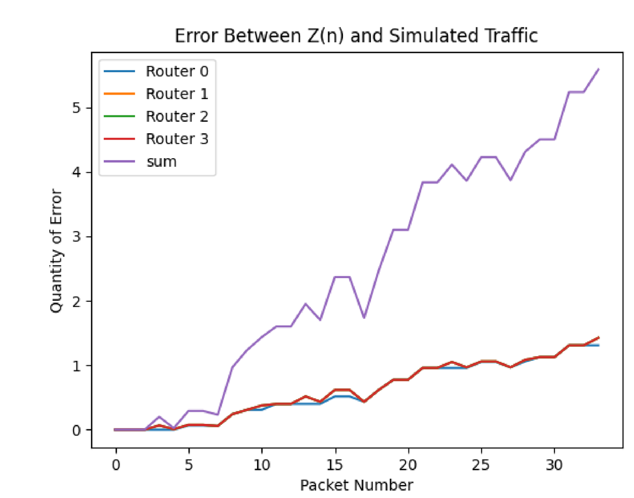


$\alpha = 2\beta$ Here we can observe that TCP connection α had half as many packets transmitted than TCP connect β in the same time.

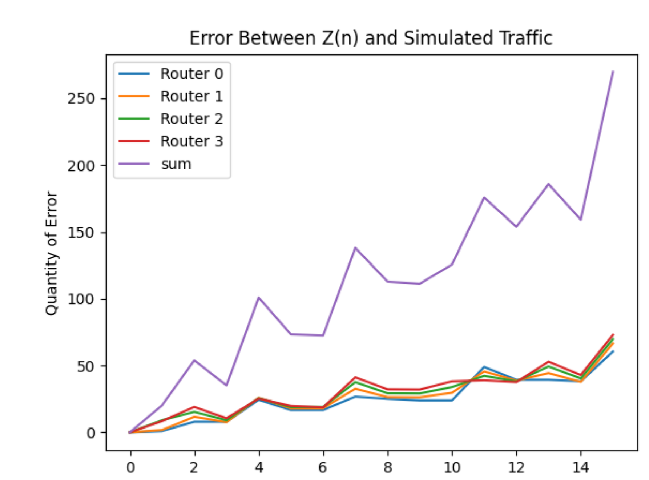


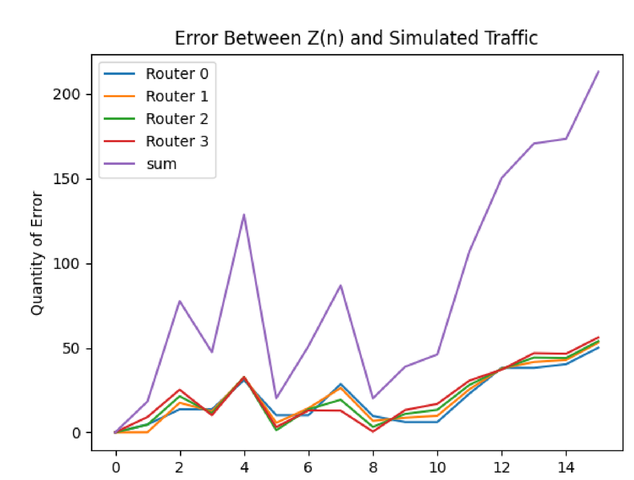
Alternating $\alpha = \frac{1}{4}\beta$ and $\alpha = \frac{1}{3}\beta$ We can observe again the quantity of packets sent for α is much larger than the quantity for β because packets for α are much smaller.





Exponential Distribution In this configuration the results had the most variance.



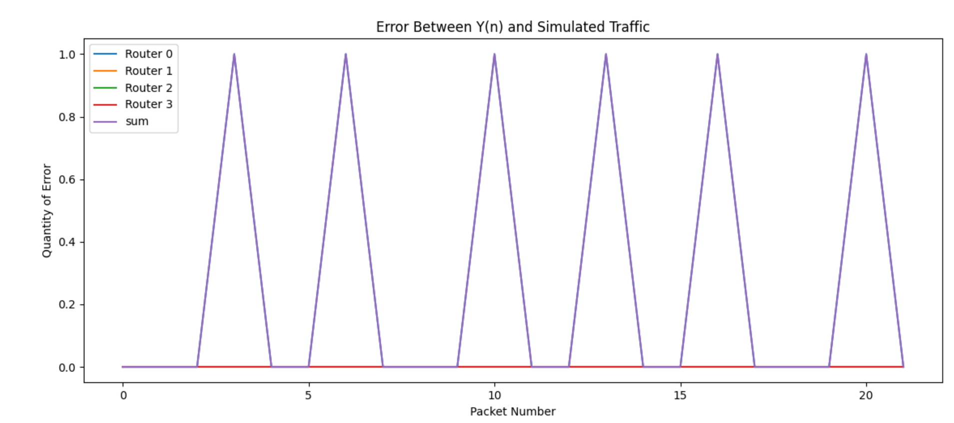


Expanding the model

Theoretical Results The graphic below demonstrates a discrepancy between the model $Y(n)$ and the simulated network traffic. This discrepancy illuminates the neglect of an initial processing time $\sigma_0(n)$ from the initial router. This can be accounted for with a initial term in $y_0(n)$ resulting in

$$y_0(n) = [y_K(n - v_{n-1}) \otimes d(n_{ack})_{K,0}] \oplus y_0(n-1) \otimes \sigma_0(n-1) \tag{4}$$

Interestingly, this margin of error in Y_0 did not impact the model $Z(n)$.This error goes away if we implement §4 in the model.



We discovered that the model requires more granularity to define the delay function $d_{i,j}$ as the delay time is a function of the packet size. Let $l_i(n)$ be the time taken to be transferred across link i, then

$$d(n)_{i,j} = \sum_{k=i}^{j} l_i(n) \otimes \sigma_i(n) \tag{5}$$

The initial construction of $y_i(n)$ works well for a single TCP connection, but neglects multiple. To expand the model to accommodate for multiple traces, let $\mathbb{J}^i$ denote the set of TCP connections running through a given switch i. Let n_j denote the nth packet for TCP connection j. For each n_j there is a unique linear path, thus $y_{i-1}(n_{j_1})$ is independent from $y_{i-1}(n_{j_2})$ $\forall j_1, j_2 \in \mathbb{J}^i$. Let $\mathbb{H}^i$ denote the unique set of packets already in the router que i at the arrival time of a packet n_{j1}. We can define the time it takes for all packets $h \in \mathbb{H}^i$ to be processed as

$$\bigoplus_{h \in \mathbb{H}^i} y_i(h)$$

Because $y_i(n_j - 1) \in \mathbb{H}^i$, we redefine the departure time of a packet n_j from a router i as

$$y_i(n_j) = [y_{i-1}(n_j) \otimes d_{i-1,i}(n_j) \bigoplus_{h \in \mathbb{H}^i} y_i(h)] \otimes \sigma_i(n_j) \tag{6}$$

This now turns the question to quantifying the elements of $\mathbb{H}$ in terms of the elements $\mathbb{J}$ for any router. Because $y_i(n_j)$ is a function of the packet size, we know that $|\mathbb{H}| \geq |\mathbb{J}|$. Furthermore, since every element $h \in \mathbb{H}$ belongs to some trace $j \in \mathbb{J}$, there exists a surjective mapping defined by the function $\Phi : \mathbb{H} \to \mathbb{J}$ from every packer $h \in \mathbb{H}$ to some trace $j \in \mathbb{J}$. If the entire topology and paths for each TCP connection are known, we know which trace j each element h belongs to just by observing the model. Defining the mapping becomes more difficult.