The Behavior of Stochastic Processes Arising in Window Protocols

Serap Savari and Emre Telatar
Lucent Technologies Bell Laboratories, Murray Hill, NJ 07974

Abstract — **Window based network flow control protocols, such as TCP, modulate the number of unacknowledged packets the protocol is allowed to have outstanding. Such protocols change the window size when they receive positive or negative acknowledgments, where the latter kind may be inferred from timeouts. Together with a communications channel that loses packets at random, such a protocol induces a stochastic process on the window size. For a broad class of window based protocols, the induced stochastic process can be analyzed using the theory of semi-Markov processes.**

I. Introduction

Consider a typical window based flow control algorithm: The algorithm is in either a *slow-start* or a *congestion avoidance* mode. In the congestion avoidance mode, the algorithm grows its window size when it detects that a packet has been successfully transmitted. The amount of the growth may depend on the current window size. If a packet loss is detected, the algorithm enters the slow-start mode, sets the window size to some small number and for each detection of a successful transmission grows the window size the growth rate may again depend on the window size. Slow-start mode ends when the window size reaches a certain threshold, at which point the algorithm enters the congestion avoidance mode. The threshold may depend, for example, on the size of the window when the packet loss was detected.

The state of such a protocol at time t can be described by specifying the mode and the window size. If the algorithm is in the slow-start mode, the threshold window size is needed to complete the state description. It is possible to summarize this information in two variables, $\vec{X}(t) = (W(t), Z(t))$. The first member of the pair $W(t)$ is the window size at time t and $Z(t)$ is an auxiliary variable: If $W(t) < Z(t)$ then the protocol is in the slow-start mode and $Z(t)$ is the threshold window size; otherwise, the protocol is in the congestion avoidance mode and $Z(t)$ does not have an operational meaning.

To complete the description of the communication system we need to model the packet losses. We will assume that the packet losses occur according to a Poisson process whose rate at a given time t depends on the window size of the protocol at that time. This allows for the packet loss rate to depend on the congestion of the link.

The evolution of $\vec{X}(\cdot)$ in time is thus governed by a stochastic differential equation:

$$\vec{X}(t+dt) = \begin{cases} \alpha(\vec{X}(t)) & \text{with prob. } \lambda(W(t))dt \\ \vec{X}(t) + \beta(\vec{X}(t))dt & \text{with prob. } 1 - \lambda(W(t))dt \end{cases}$$

where $\lambda(W(t))$ is the packet loss detection rate when the state of the process is $W(t)$. Ordinarily, one will have an $\alpha(\cdot)$ such that the window size is reduced when a packet loss is detected, and a $\beta(\cdot)$ which increases the window size when an acknowledgment is received.

The formulation above is very general. Here we will focus only on models in which the auxiliary variable Z stays constant during a period where no packet loss is detected; i.e., we will only consider functions β whose second component is identically zero. Even with this restriction this formulation allows us to accommodate a larger class of protocols and a larger class of packet-loss models than formulations considered before.

Processes of this type lend themselves to an analysis based on the theory of semi-Markov processes. Even though the state process $\vec{X}(t)$ is Markovian, the window size $W(t)$ is, in general, not. Nonetheless, $W(t)$ is a semi-Markov process. Namely, we can identify epochs $S_0, S_1, \ldots$, such that $W_n = W(S_n^+)$ is a discrete time Markov process, and the transition times $T_n = S_{n+1} - S_n$ are random, but depend only on W_n and W_{n+1}. In particular, we take S_n to be the start of the nth congestion avoidance phase.

II. Analysis

We refer the reader to [1] for the details of the analysis for various choices of α and β. Here we summarize the key ideas in this analysis, by focusing on a protocol without a slow-start mode. This case is the basis of the analyses of the protocols with slow-start that are considered in [1]. In this protocol $Z(t)$ is identically zero. S_n is now the time of the nth packet loss.

Since between packet losses the evolution of the process is deterministic, it is sufficient to analyze the properties of W_n to capture the behavior of the process at all times. Accordingly we proceed by computing the transition probability density of the chain and finding its steady state distribution. We then compute the statistics of the continuous time process from the statistics of the discrete time chain.

As an example, suppose the window size is halved at each packet loss, and is grown according to $W'(t) = \beta/W(t)$ in between packet losses. If the loss rate is $\lambda W(t)$ at time t, then W_n has density

$$\pi(w) = \frac{\lambda w^2}{\beta} \sum_{k=0}^{\infty} a_k \exp\Bigl(-8^{k+1} \frac{\lambda w^3}{3\beta}\Bigr)$$

with a_k given by the recursion

$$a_0 = \frac{8}{\prod_{k>0} 1 - 8^{-k}}, \quad a_k = a_{k-1} \frac{8}{8^k - 1}.$$

The example above is clearly simplistic in various ways. The real protocols do have slow-start phases. This poses only minor technical difficulties, and a similar analysis with more complicated looking results can be carried out. The novelty of the model presented here lies more in its effort to capture the effect of the window size on the loss rate through $\lambda(W(t))$. This only takes into account the effect of one user's window size on the loss rate, and is thus appropriate in communication links with a dominant user. In principle the analysis here can be extended to cases where there are two dominant users; this will require the analysis of a two-dimensional process.

References

[1] S. Savari and E. Telatar, "The Behavior of Certain Stochastic Processes Arising in Window Protocols," Submitted to IEEE Transactions on Information Theory, 1998.