Robust Stability Analysis of a class of Smith Predictor-based Congestion Control algorithms for Computer Networks

Luca De Cicco, Saverio Mascolo and Silviu I. Niculescu

Abstract—Congestion control is a fundamental building block in packet switching networks such as the Internet due to the sharing of communication resources. It has been shown that the plant dynamics is essentially made of an integrator plus time delay and that the a proportional controller plus a Smith predictor is a simple and effective controller. It has been also shown that the today running TCP congestion control can me modeled using a Smith predictor plus a proportional controller [8]. Due to the importance of this control structure in the field of data network congestion control, we analyze the robust stability of this controller in the face of delay uncertainties that in data networks are present due to queuing. In particular, by applying a simple geometric approach [3], we derive a bound on the proportional controller gain which is necessary and sufficient to guarantee stability given a bound on delay uncertainty.

I. Introduction

Time delays are often present in feedback control systems due to reasons such as the transport of material or information. From the control theoretic point of view it is well-known that an increase of the time delay may lead to instability of the closed loop system and to performance degradation as well.

The Smith principle is a classic approach which is often employed to design controllers for time delay systems [15]. It is known that, by assuming exact knowledge of both the plant model and time delay, controllers designed using a Smith predictor are very effective in counteracting the effect of time delays. Robustness of the Smith predictor with respect to uncertainties in the knowledge of the time delay has been extensively studied since 1980 [13], [17].

The Internet represents a relevant example of time delay system due to the presence of delays that are caused by the propagation of the information, which is sent in form of data packets, from a source to a destination through a series of communication links and router queues.

A cornerstone component of the Internet protocol stack is the end-to-end congestion control which has been implemented in the TCP in the late '80s by V. Jacobson in order to avoid congestion and preserve network stability [6]. Several fluid models have been proposed for the TCP congestion control algorithm in order to analytically study the stability of the network under different scenarios [4], [8] .

Luca De Cicco is a PhD student of Dipartimento di Elettrotecnica ed Elettronica, Politecnico di Bari, Via Orabona 4, Italy ldecicco@poliba.it

Saverio Mascolo is a Faculty of Dipartimento di Elettrotecnica ed Elettronica, Politecnico di Bari, Via Orabona 4, Italy mascolo@poliba.it

Silviu Iulian Niculescu is a Faculty of Laboratoire de Signaux et Systèmes (L2S), Supélec, 3, rue Joliot Curie, 91190 Gif-sur-Yvette, France Silviu.Niculescu@lss.supelec.fr

In [8] a simple model of the plant made of an integrator (modelling the bottleneck queue) plus two time delays (modelling forward and backward delays), has been proposed along with a Smith predictor plus a proportional controller. The paper also shows that the Smith predictor controller with a proportional gain models the congestion control law which is employed in the today running TCP congestion control algorithm. Moreover, the model presented in [8] has been employed to design and implement a rate-based congestion control algorithm which is TCP friendly [2].

Measurement of the plant time delay to be used in the Smith predictor can be affected by uncertainties due to the fact that the time delay is made of a constant propagation delay plus a time-varying queueing delays. To the purpose, the standard TCP [14] estimates the Round Trip Time (RTT) through time-stamping in order to set the retransmission timeout (RTO) [14] which is needed for detecting heavy congestion episodes in the network. The RTT is defined as the time that elapses from when a segment is sent until the corresponding acknowledgement segment is received by the sender. In the standard TCP implementation, the RTT is measured each RTT seconds, whereas no measurements are taken on retransmitted segments due to the Karn's algorithm [7] in order to avoid spurious timeouts. For these reasons TCP, in its standard implementation, does not provide an accurate measure of RTT. In order to overcome this issue another optional scheme has been proposed and standardized in [5] which makes use of timestamps in an optional field of the TCP header. However, even if the timestamp option is employed by both peers of the communication, the granularity chosen for TCP timestamps is implementationdependent. In a recent work an extensive measurement campaign on RTTs as been carried out [16]. Authors used 500 servers and found that 76\% of the servers had timestamping option enabled, and out of these servers 37% used a 100 ms granularity, 55% a 10 ms granularity and only 7% of them had a granularity of 1 ms.

A preliminary study on robust stability of a proportional Smith predictor used for congestion control in data networks has been carried out by using the Nyquist criterion in [9]. It revealed that in order to guarantee asymptotic stability it is sufficient that $\Delta < 1/k$ where Δ represents the delay uncertainty and k is the gain of the proportional controller.

The goal of this paper is to provide a characterization of the robust stability of system introduced in [8] by applying the geometric approach which has been developed in [3].

The rest of the paper is organized as follows: in Section II we briefly review the model [8] of the closed loop congestion

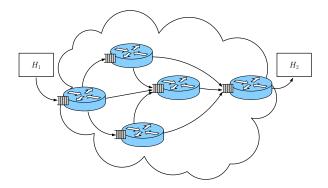


Fig. 1. Network connection between the hosts H_1 and H_2

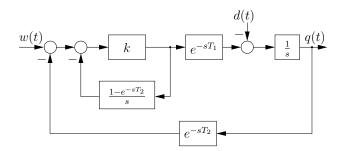


Fig. 2. Functional block of the congestion control model

in a general packet switching networks; in Section III we apply the geometrical approach [11] in order to find the stability crossing curves of the system; in Section IV we present the robustness analysis; finally Section V concludes the paper.

II. CONGESTION CONTROL MODEL

A network connection is basically made by a set of communication links and store-and-forward nodes (routers) where packets are enqueued before being routed to the destination (see Figure 1). Congestion can arise when packets arrive at a rate which is above the capacity of the output link so that the router queue builds up until it is full and it starts to drop packets.

In [8] a model the Internet flow and congestion control as a time delay system is provided and it is shown that different variants of TCP congestion control algorithms can be modeled in a unified framework by proper input shaping of the proportional Smith predictor controller [10]. In particular, the model consists of a feedback loop in which two time delays are present as it is shown in Figure 2: T_1 models the propagation time of a packet from source to the bottleneck and T_2 models the propagation time from the bottleneck to the destination and then back to the sender. The round trip time of the connection is $T = T_1 + T_2$.

The bottleneck queue is modeled by a simple integrator 1/s and the controller is a proportional Smith predictor with gain K. The reason for using a simple proportional controller is that in this way the closed-loop dynamics can be made that of a first-order system with time constant 1/k delayed by T_1 . Thus, the step response of the system can be made faster by increasing the proportional gain k providing an always

stable system without oscillations or overshoots. Moreover this choice provides a controller in which only one design parameter, i.e. the gain k has to be tuned.

Model mismatch are known to affect the closed loop dynamics when a Smith predictor controller is employed. In this case, it is worth noting that the only source of mismatch between the model and the actual plant is the entity of the delay (see Section I) whereas the model of the bottleneck queue is an integrator. In the next sections we will give simple rules in order to tune the design parameter k in order to retain asymptotic stability when the measure of time delay T is uncertain.

Finally we remark that a Smith predictor controller when designing a congestion control algorithm for data networks is recommended since using PID controllers would provide an unacceptable sluggish system due to large delays involved in communication networks [1].

III. STABILITY CROSSING CURVES IN THE PARAMETERS SPACE

A. Review of the geometrical approach

We start by briefly reviewing the geometrical approach which we will employ to analyze the robust stability of the considered system [11]. The reader is advised to refer to [3] for a complete description of the method. We denote with $a(s;\tau_1,\tau_2)$ the characteristic function of the closed-loop system where τ_1 represents the nominal delay used in the Smith predictor and $\tau_2=\tau_1+\Delta$ represent the actual plant delay delay affected by a mismatch Δ . It is easy to show that the characteristic function in this case is given by:

$$a(s; \tau_1, \tau_2) = 1 - h(s)e^{-\tau_1 s} + h(s)e^{-\tau_2 s}$$
 (1)

where h(s) is the transfer function of the closed loop system when no delays are present in the loop:

$$h(s) = \frac{C(s)G_0(s)}{1 + C(s)G_0(s)}$$

with $G_0(s)$ the delay free plant and with C(s) the controller transfer function.

In order to analyze the stability of the system we look for the solutions of the characteristic equation:

$$a(j\omega; \tau_1, \tau_2) = 0 \tag{2}$$

In this way we are able to find all the conditions under which the system has at least one pole on the imaginary axis. The geometrical approach relies on the observation that the three terms of the characteristic function (1) can be seen as vectors in the complex plane. The equality $a(s;\tau_1,\tau_2)=0$ can be represented in the complex plane via an isosceles triangle as it is shown in Figure 3. Thus, equation (2) is equivalent to the following conditions:

1) The triangular inequality must hold for the triangle shown in Figure 3, which implies that:

$$|h(j\omega)| \ge \frac{1}{2} \tag{3}$$

2) Equation (2) must satisfy the phase rule;

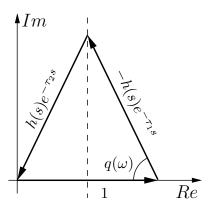


Fig. 3. Triangle formed by the three vectors when the characteristic equation holds

3) The sum of the internal angles of the isosceles triangle must be equal to π ;

The solution of (3), which does not depend on time delays τ_1 or τ_2 , forms the *frequency crossing set* Ω which is the union of a finite number of intervals of finite length $\Omega_1,\Omega_2,\ldots,\Omega_N$ [3]. For any $\omega>0$ which belongs to the frequency crossing set there exists at least a pair (τ_1,τ_2) in the parameters space such that the system has at least one imaginary pole. The conditions 2 and 3 imply that for all $\omega\in\Omega$ all the couples $(\tau_1,\tau_2)\in\mathbb{R}^2_+$ which satisfy $a(j\omega;\tau_1,\tau_2)=0$ can be found using the following equations:

$$\tau_1^{u\pm} = \frac{\angle h(j\omega) + 2u\pi \pm q}{\omega} \tag{4}$$

$$\tau_2^{v\pm} = \frac{\angle h(j\omega) + (2v-1)\pi \mp q}{\omega} \tag{5}$$

where u and v are integers such that the corresponding $\tau_1^{u\pm}$, $\tau_2^{v\pm}$ are non negative and q represents the internal angle of the isosceles triangle:

$$q(\omega) = \arccos\left(\frac{1}{2|h(j\omega)|}\right)$$

In order to understand the meaning of equations (4) and (5) let us fix $u=\overline{u}$ and $v=\overline{v}$: if ω varies in Ω_i and we evaluate (4) and (5) for both positive and negative signs we obtain two curves in the parameter space (τ_1,τ_2) which we denote $\mathcal{T}^{i+}_{\overline{u},\overline{v}}$ and $\mathcal{T}^{i-}_{\overline{u},\overline{v}}$ respectively. It is worth noting that the curves $\mathcal{T}^{i}_{u,v}=\mathcal{T}^{i-}_{u,v}\cup\mathcal{T}^{i+}_{u,v}$ can be either open curves or closed curves depending on the set Ω_i we are considering. In particular, if the left end of Ω_i is 0 then the associated curve is an *open curve* with both ends approaching ∞ when $\omega \to 0$. On the other hand, if the left end of Ω_i is not 0 then $\mathcal{T}^i_{\overline{u},\overline{v}}$ is a *closed curve* [3].

We define the *stability crossing curves* \mathcal{T} in the τ_1, τ_2 plane as the union of all the curves $\mathcal{T}_{u,v}^i$ when $i \in \{1, \ldots, N\}$, and u and v vary in the set of integers.

Finally, it is important to point out that when we cross a stability crossing curve in the τ_1, τ_2 plane, two poles cross the imaginary axis on the complex plane [3].

B. Stability crossing curves of the computer network congestion control model

In order to characterize the impact of the delay uncertainty on the stability of the considered feedback system we apply the geometric approach we have reviewed in Section III-A. We suppose that the system described in Section II is affected by a delay uncertainty Δ which is bounded by $\delta>0,$ i.e. $|\Delta|<\delta$. By considering the delay uncertainty, the characteristic equation of the system can be rewritten as follows:

$$1 + \frac{k}{s} - \frac{k}{s}e^{-\tau_1 \cdot s}(1 - e^{-\Delta \cdot s}) = 0 \tag{6}$$

where τ_1 represents the nominal round trip time (RTT) of the considered connection, which is used in the Smith predictor, and $\tau_2 = \tau_1 + \Delta$ is the actual plant time delay.

By multiplying by s/(s+k) both sides of (6) we obtain:

$$1 - \frac{k}{s+k}e^{-\tau_1 s} + \frac{k}{s+k}e^{-(\Delta+\tau_1)s} = 0 \tag{7}$$

which is in the form of (1). We are interested in characterizing the stability of the system when τ_1 , τ_2 and k vary in \mathbb{R}_+ .

By making the change of variable z = s/k we obtain:

$$1 - \frac{1}{z+1}e^{-h_1z} + \frac{1}{z+1}e^{-h_2z} = 0$$
 (8)

where $h_1 = k\tau_1$ and $h_2 = k\tau_2$, which reduces the free parameters to two. It is worth to notice that the transformation from (7) to (8) simply involves a *scaling* of the closed-loop eigenvalues by 1/k, thus indicating a natural trade-off between gain and delay since when k increases the closed loop poles approach to the imaginary axis [12].

We are now ready to study the stability of the original system in the h_1, h_2 plane regardless the value of the proportional gain k.

First of all by applying (3) we find that the crossing set is made by the single interval $\Omega=[0,\sqrt{3}]$ which means that the stability crossing curves in the h_1 , h_2 plane are open curves which extend to infinity when $\omega\to 0$. By using (4) and (5) the stability crossing curves of the considered system result the following:

$$h_1^{u\pm}(\omega) = \frac{-\arctan\omega + 2u\pi \pm \arccos\left(\frac{\sqrt{1+\omega^2}}{2}\right)}{\omega}$$
 (9)

$$h_2^{v\pm}(\omega) = \frac{-\arctan\omega + (2v-1)\pi \mp \arccos\left(\frac{\sqrt{1+\omega^2}}{2}\right)}{\omega}$$
(10)

Figure 4 shows the stability crossing curves of the considered system. We start by considering the h_2 axis $(h_1=0)$, which means that we are employing a simple proportional controller without the Smith predictor. By starting from the origin and increasing the value of h_2 the first curve is crossed at $h_2=\pi/2$ which means that the system becomes unstable for $h_2>\pi/2$. On the other hand, the axis h_1 represents the system in which no delay affects the plant, but the Smith predictor is in the controller. Figure 4 shows that the system is stable for all the delays in the Smith predictor.

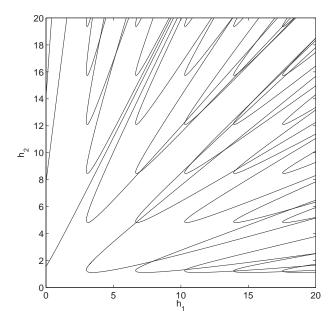


Fig. 4. Stability crossing curves for the considered system

Finally, points on the positive bisector represent the case of perfect matching of nominal delay τ_1 with the actual delay τ_2 . Indeed, if we move on this line no curves will be crossed since the Smith predictor in this case provides a stable system regardless the value of the proportional gain k.

IV. ROBUST STABILITY ANALYSIS

In this section we will develop an analysis of the robust stability of the considered system by using the stability crossing curves we have shown in the previous Section. We already know that the considered system is always asymptotically stable for any delay τ_1 and any proportional gain k as far as the delay uncertainty is zero thanks to the perfect compensation of the time delay τ_1 provided by the Smith predictor. In the h_1, h_2 plane this condition means that the system is asymptotically stable on the positive bisector.

In order to characterize the robustness of the system in the face of delay uncertainties we compute the maximum delay mismatch which still preserve stability. Thus, the problem here is to look for the maximum deviation δ with respect to a generic point (τ_1^*, τ_1^*) with $\tau_1^* \geq 0$ which lies on the positive bisector such that the system is stable for any (τ_1, τ_2) which satisfies:

$$|\tau_2 - \tau_1^*| < \delta$$

We remark that solving the maximum admissible delay uncertainty problem is equivalent to find the minimum distance between the stability crossing curves and a generic point on the positive bisector of the h_1, h_2 plane.

Thus for any $\tau_1^* > 0$ we have to solve:

$$\delta(\tau^*) = \min_{u,v} \min_{\tau_2^{v^{\pm}} \in \mathcal{T}} |\tau_2^{v^{\pm}} - \tau_1^*|$$
 (11)

so that the maximum delay to retain stability is:

$$\delta = \min_{\tau^* \in \mathbb{R}_+} \delta(\tau_1^*) \tag{12}$$

Proposition 1: A necessary and sufficient condition for the asymptotic stability of the system regardless the value of the nominal delay τ_1 is:

$$|\Delta| < \frac{\alpha}{k} \tag{13}$$

where Δ is the delay uncertainty, $\alpha = 1.4775$ and k is the proportional gain of the controller.

Proof: We start by considering the stability crossing curves in the parameters space h_1 , h_2 . In order to find the minimum distance between the stability crossing curves and a generic point of positive bisector of the h_1 , h_2 plane we evaluate the tangent to the crossing curves with direction parallel to the positive bisector:

$$\frac{dh_2}{dh_1} = 1 \Leftrightarrow \frac{dh_2}{d\omega} \frac{d\omega}{dh_1} = 1 \Leftrightarrow \frac{dh_1}{d\omega} = \frac{dh_2}{d\omega}$$
 (14)

To the purpose we look for a subset \overline{T} of the stability crossing curves T that are the "closest" curves to the positive bisector. By considering a generic curve $T_{u,v}$ and by applying (9) and (10) it turns out that for all u and v and for all $u \in \Omega$ it holds $h_2^{v+} - h_1^{u+} < h_2^{v-} - h_1^{u-}$ so that it is sufficient to consider only the curves $T_{u,v}^+$ in the region $h_2 > h_1$ and the curves $T_{u,v}^-$ in the region $h_2 < h_1$, since they will be the closest ones to the positive bisector. Thus, we can refer without loss of generality to the generic curve of T as $T_{u,u+i}$ for all i and u in the integers. Straightforward computations on (9) and (10) give:

$$h_2^{u+i} - h_1^u > h_2^{u+i-1} - h_1^u$$

which means that when i decreases the curves $\mathcal{T}_{u,u+i}$ will move downwards in the h_1,h_2 plane. Figure 5 shows the values of u and v for the curves $\mathcal{T}_{u,u}$ and $\mathcal{T}_{u,u+1}$. It is then easy to show that if we set v=u we obtain the closest curves to the positive bisector in the region $h_2 < h_1$ whereas the curves with v=u+1 are those which are closest to the positive bisector in the region $h_2 > h_1$. In conclusion we can restrict our search to the set:

$$\overline{\mathcal{T}} = \mathcal{T}_{u,u}^- \cup \mathcal{T}_{u,u+1}^+$$

for all u in the integers. Let us consider the region $h_2 > h_1$ i.e. we consider the subset $\mathcal{T}_{u,u+1}^+$. By considering (14) after straightforward computations we get the following equation:

$$\arccos\left(\frac{\sqrt{\omega^2+1}}{2}\right) + \frac{\omega^2}{\sqrt{\omega^2+1}\sqrt{3-\omega^2}} - \pi(v-u-\frac{1}{2}) = 0$$
(15)

with $\omega \in \Omega$. By letting v=u+1 the equation (15) has the unique solution $\overline{\omega}=1.3483$ rad/s in Ω which is independent of u. If we substitute this value in (9) and (10) we obtain:

$$h_1(\overline{\omega}) = h_1 = 4.6601u - 0.2654$$

$$h_2(\overline{\omega}) = h_2 = 4.6601v - 3.4480$$

Thus, all the points belonging to the the curves $\mathcal{T}_{u,u+1}^+$ having a tangent which is parallel to the positive bisector, lie on the line:

$$h_2 = h_1 + 1.4775 \tag{16}$$

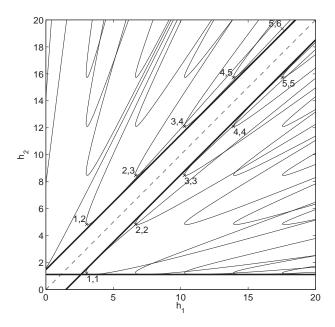


Fig. 5. Stability regions for the considered system

For this reason we can conclude that the maximum uncertainty, in the h_1, h_2 coordinates is 1.4775. The proof is completed by recalling that $h_1 = k\tau_1$ and $h_2 = k\tau_2$ and that $\tau_2 = \tau_1 + \Delta$. Thus, we finally obtain:

$$h_2 - h_1 < 1.4775 \Rightarrow k\Delta < 1.4775 \Rightarrow \Delta < \frac{1.4775}{k}$$
 (17)

It is worth to notice that the same procedure can be followed in the case v = u which leads to the inequality:

the case
$$v=u$$
 which leads to the frequency.
$$h_1 - h_2 < 1.4775 \Rightarrow -k\Delta > 1.4775 \Rightarrow \Delta > -\frac{1.4775}{k}$$
 (18)

Thus, by considering both (17) and (18) we can conclude that:

$$|\Delta| < \frac{1.4775}{k} \tag{19}$$

In order to prove the necessity of the condition (19) let us consider the curves $\mathcal{T}_{u,u+1}$. The points of the curve $\mathcal{T}_{u,u+1}$ that correspond to the frequency $\omega=1.3483\,\mathrm{rad/s}$ lie on the line described by (16) so that the maximum delay uncertainty admissible for those points is exactly α/k . If we select a larger value for δ the system will become unstable at least on those points. This conclude the proof.

Remark 1: The fact that the maximum uncertainty allowed does not depend on the nominal delay τ_1 is a very nice feature of the Smith predictor based controller. This makes the controller effective even with large delays.

Remark 2: The condition (13) expresses a trade-off between the maximum delay mismatch δ and the proportional gain that can be used to tune the controller gain k.

Remark 3: This result improves the robust stability condition $|\Delta| < 1/k$ found in [9].

Proposition 2: The system is stable, regardless the value of τ_1 , if the delay uncertainty Δ satisfies the following

inequality:

$$-\tau_1 < \Delta < -\tau_1 + \frac{\beta}{k} \tag{20}$$

with $\beta = 1.1188$.

Proof: The proof follows the same arguments of Proposition 1, therefore it is omitted.

Remark 4: The condition (20) implicitly requires the delay uncertainty Δ to be negative, that is, the nominal delay τ_1 should be always below the actual delay of the plant τ_2 . For this reason condition (20) it has no particular meaning for the characterization of controller robustness, since the sign of the uncertainty is not known a priori.

V. CONCLUSIONS

In this paper we have analyzed the robust stability of a very important class of congestion control algorithms when delay uncertainties are present. We have shown how the geometrical approach developed in [11] can be easily applied in order to find stability bounds on the parameter of the controller. Moreover, we found a very easy to use necessary and sufficient condition on the gain of the proportional controller k in order to retain asymptotic stability regardless the value of the nominal delay τ_1 . Such a result suggests that congestion control algorithms that employ controllers made by a Smith predictor plus a proportional gain can be easily tuned in order to be robust to delay uncertainty.

REFERENCES

- K.J. Astrom, "PID Controllers: Theory, Design, and Tuning", Instrument Society of America, 1995
- [2] L.A. Grieco and S. Mascolo, "Adaptive Rate Control for Streaming Flows over the Internet", ACM Multimedia Systems Journal, Jun. 2004, Volume 9, Issue 6, pp. 517 - 532
- [3] K. Gu, S.-I. Niculescu and J. Chen, "On stability crossing curves for general systems with two delays", Journal of Mathematical Analysis and Applications, Vol. 311, Issue 1, 2005, pp. 231-253
- [4] C. V. Hollot, V. Misra, D. Towsley and G. Wei-Bo, "Analysis and design of controllers for AQM Routers supporting TCP flows". IEEE Transactions on Automatic Control, 2002, 47(6), pp. 945-959
- [5] V. Jacobson, R. Braden and D. Borman, "TCP Extensions for High Performance", RFC 1323, May 1992
- [6] V. Jacobson, "Congestion Avoidance and Control", Proc. of SIG-COMM '88 in ACM Computer and Communication Review, vol. 18, no. 4, pp. 314-329
- [7] P. Karn and C. Partridge, "Estimating Round-Trip Times in Reliable Transport Protocols", Proc. SIGCOMM '87, Stowe, VT, August 1987.
- [8] S. Mascolo, "Congestion control in high-speed communication networks using the Smith principle Automatica", vol. 35, no. 12, Dec. 1999, pp. 1921-1935. Special Issue on "Control methods for communication networks
- [9] S. Mascolo, "Modeling the Internet congestion control as a time delay system: a robust stability analysis", IFAC Workshop on Time-delay Systems, September 8 - 10, 2003
- [10] Saverio Mascolo, "Modeling the Internet congestion control using a Smith controller with input shaping", Control Engineering Practice, April 2006, Volume 14, Issue 4, pp. 425-435, Elsevier
- [11] C.-I. Morarescu, S.-I. Niculescu and K. Gu, "On the geometry of stability regions of Smith predictors subject to delay uncertainty", IMA Journal on Mathematical Control and Information, 2006
- [12] S.-I. Niculescu and W. Michiels, "Stabilizing a Chain of Integrators Using Multiple Delays", IEEE Transactions on Automatic Control, 2004, vol. 49, pp. 802-806
- [13] Z. J. Palmor, "Stability properties of Smith deat-time compensator controllers", Int. J. Control, 32, pp. 937-949, 1980
- [14] J. Postel, "Transmission Control Protocol", RFC 793, Sept 1981
- [15] O. J. M. Smith, "A controller to overcome dead time", Instrum. Soc. Am. J., vol. 6, pp 28-33

- [16] B. Veal, K. Li and D. Lowenthal, "New Methods for Passive Estimation of TCP Round-Trip Times", in Proc. Passive and Active Measurements (PAM), 2005
- Measurements (PAM), 2005

 [17] K. Yamanaka and E. Shimemura, "Effects of mismatched Smith controller on stability in systems with time-delay", Automatica, vol. 23, pp. 787-791, 1987