Available online at www.sciencedirect.com
ScienceDirect
Automatica 44 (2008) 1402–1410
automatica
www.elsevier.com/locate/automatica

Brief paper

A neuro-adaptive congestion control scheme for round trip regulation☆

Christos N. Houmkozlis, George A. Rovithakis*

Aristotle University of Thessaloniki, Department of Electrical and Computer Engineering, Thessaloniki 54124, Greece

Received 22 December 2006; received in revised form 27 September 2007; accepted 9 October 2007
Available online 21 February 2008

Abstract

A novel neuro-adaptive congestion controller is presented, capable of regulating the per packet round trip time (RTT) around a piecewise constant desired RTT, thus achieving almost piecewise constant delay. The controller is implemented at the source and is proven robust against modeling imperfections, exogenous disturbances (UDP traffic) and delays (propagation, queueing). The notion of communication channels is introduced for throughput improvement. The analysis is nonlinear and the tools used are approximation-based control and linear-in-the-weights neural networks. The proposed controller is guaranteed to be saturated. Moreover, modifications are also provided to achieve rate reduction whenever congestion is detected. Simulation studies illustrate the performance of the proposed control scheme and compare it with other well-established congestion control mechanisms.
© 2008 Elsevier Ltd. All rights reserved.

Keywords: Congestion control; Neural adaptive control; Delay regulation

1. Introduction

A dominant problem in the design of packet-switching networks is the development of congestion control algorithms. The congestion control mechanisms in the Internet consist of a source algorithm that dynamically adjusts source rates based on congestion in their paths, and a link algorithm that updates a certain congestion measure at each link and feeds it back, implicitly or explicitly, to the sources that use the link. Initially some heuristic methods were developed despite their wide applicability, simulation studies (Low, Paganini, & Doyle, 2002) have demonstrated that they may be ill-suited for networks where both communication delay and network capacity can be large (Wei, Jin, Low, & Hegde, 2006). This has motivated research on theoretical understanding of TCP congestion control and the search for protocols that scale properly so as to maintain stability in the presence of these variations. In particular, Kelly, Maulloo, and Tan (1998) and Low and Lapsley (1999) have developed an optimization-based framework that provides an interpretation of various congestion control mechanisms. They proposed a primal algorithm for source algorithm and a dual algorithm for the link algorithm, which generalize the Additive Increase/Multiplicative Decrease (AIMD) congestion avoidance strategy (Jacobson, 1988) to large-scale networks.

Moreover, the introduction of new types of services in the Internet has underlined the importance of Quality of Service (QoS) even in congested network conditions. This has further motivated researchers to develop new mechanisms capable of guaranteeing Quality criteria (Cai, Shen, Mark, & Pan, 2006; Weber & Veciana, 2005), while preventing network from congestion collapse and throughput from starvation. Such applications typically demand higher sensitivity to time constraints (i.e., delay), so as to cause the least possible degradation of quality of service (QoS) to the underlying users, silently implying the existence of a control algorithm capable of regulating the round trip time, avoiding either overflow or empty queues.

Current control theoretic approaches are mainly focused on designing link algorithms, implemented as congestion controllers in the routers (Kim, 2006; Tan, Zhang, Peng, & Chen, 2006). System stability is analyzed under the assumption that all sources transmit packets following a certain fluid flow model that incorporates the AIMD mode of TCP. In

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Derong Liu under the direction of Editor Miroslav Krstic.
* Corresponding author.
E-mail addresses: houm@eng.auth.gr (C.N. Houmkozlis), robi@eng.auth.gr (G.A. Rovithakis).

0005-1098/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2007.10.012

fact, local stability around an equilibrium is addressed, since linearized versions of the fluid flow model and queue dynamics are considered. From a systems theoretical perspective, these link algorithms constitute the controllers, while the rest of the network is considered as the plant. A necessary requirement — characteristic of a candidate link algorithm to be implementable, should be its simplicity. In this way the total number of computations performed owing to congestion control in a router is kept to a minimum, a property of great significance especially when high-speed networks are considered. All the aforementioned techniques propose linear congestion controllers owing to system linearization. Thus, they fulfill the design constraint of simplicity. On the other hand, the intrinsic highly nonlinear and uncertain character of a packet-switching network is lost.

An alternative approach, is to design congestion controllers to operate in the source (Alpcan & Basar, 2005; Floyd, Handley, Padhye, & Widmer, 2000; Kunniyur & Srikant, 2003). In this respect, source controls the packet-transmission rate, with the plant now being the network as seen from the source. Such schemes though possess certain critical issues that should be carefully addressed. Specifically: (1) A model for the plant (network) has to be devised, otherwise stability analysis is impossible, (2) Determining a desired and feasible system performance is certainly a non-trivial task, owing to the highly uncertain and dynamically varying nature of the Internet, (3)The controlled system should not be only stable, but additionally robustly stable against modeling imperfections, exogenous disturbances (e.g., UDP traffic) and delay and (4) The output of any proposed control scheme should be saturated.

Fortunately, in the source control theoretic approaches the simplicity constraint is now relaxed. Moreover, the recently developed systems discipline of approximation-based control (i.e., neural network control, fuzzy control etc.) that was originally proposed for on-line black box system identification and robust adaptive control of highly uncertain and nonlinear dynamical systems, has now reached a maturity level, making it a promising working framework for developing source implemented congestion controllers. Along these lines Haeri and Rad (2006) used linear recursive system identification to model the round trip dynamics and consequently an adaptive TCP congestion control strategy is proposed. The control signals determines the window size in the transmission, such that the plant output (average round trip) tracks a desired average round trip, which is estimated on-line using a mechanism similar to the one TCP Vegas employs, to estimate the queue level in the routers. Besides the linear model used, no robustness analysis is performed to guarantee system stability against modeling errors, exogenous disturbances and delays, thus leaving certain critical issues open for future investigations.

A different scheme was proposed in Houmkozlis and Rovithakis (2005) involving a nonlinear neural network model to capture the round trip dynamics. A nonlinear adaptive controller to determine the per packet sending rate was developed, to track an on-line estimated desired per packet round trip. The congestion control scheme was proven robust against modeling imperfections. However, the no delay, no disturbance (uncontrollable traffic) assumptions, restrict severely the applicability.

In this paper we extend our previous work in the direction of relaxing the no delay, no uncontrollable traffic assumptions, while addressing the aforementioned critical issues (1)–(4). The proposed source control theoretic framework is comprised of three dominant modules: (i) an on-line desired round trip estimator, whose main objective is to output reliable piecewise constant predictions of feasible (i.e., achievable) round trip times; (ii) an adaptive and saturated transmission rate controller, capable of precise tracking the desired round trip time commands that are downloaded from the first module and (iii) a future path congestion level estimator, which provides critical information to the aforementioned two modules. To improve throughput, the notion of parallel operating communication channels (virtual sources) is introduced.

The analysis is nonlinear and the tools used are approximation-based control and linear-in-the-weights neural networks. The proposed congestion control scheme guarantees the uniform ultimate boundedness of the tracking error with respect to an arbitrarily small neighborhood of zero, plus the uniform boundedness of all other signals in the closed loop. Moreover, our controller is proven robust against modeling imperfections, exogenous disturbances (i.e., UDP traffic) and delays. Besides guaranteeing the saturated character of the proposed rate controller, an extra modification is also provided to further guarantee rate reduction whenever congestion is detected. Regulating round trip time to follow a piecewise constant profile, is equivalent to queueing delay and queue length regulation around a piecewise constant profile, thus introducing certain QoS characteristics. Furthermore, it is expected that the network operates in a low congestion, high throughput regime. Simulation studies have been performed, to illustrate the attributes of the proposed control scheme and to compare it with other well-established congestion control mechanisms.

The paper is organized as follows. Section 2 presents the control problem and preliminaries. Section 3, analyzes the proposed neural network rate control algorithm, as well as the future path congestion estimator. The group sending time estimation algorithm is presented in Section 4, while illustrative simulation studies are performed in Section 5. Finally, we conclude in Section 6.

2. Problem formulation and preliminaries

2.1. Packet-switching network system

We consider a general packet-switching network. The network topology is characterized by a set of sources/receivers $\mathcal{C} = \{1, 2, \ldots, n\}$, a set of nodes $\mathcal{Q} = \{1, 2, \ldots, m\}$ and a set of links $\mathcal{L} = \{1, 2, \ldots, l\}$ connecting the nodes. Each link $l \in \mathcal{L}$ has an associated queue with maximum capacity B_l. A source $S \in \mathcal{C}$ has to transmit an application of prespecified amount of packets N to a destination $D \in \mathcal{C}$ through the network whose transmission rate (u) is controlled by an appropriately designed

Download English Version:

https://daneshyari.com/en/article/698373

Download Persian Version:

https://daneshyari.com/article/698373

Daneshyari.com