Automatica 48 (2012) 1017–1030

Contents lists available at SciVerse ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Stochastic optimal control of unknown linear networked control system in the presence of random delays and packet losses[☆]

Hao Xu[a,1], S. Jagannathan[a], F.L. Lewis[b]

[a] Electrical and Computer Engineering Department, Missouri University of Science and Technology, Rolla, MO 65401, USA

[b] Electrical Engineering Department, AARI, University of Texas at Arlington, Arlington, TX 76118, USA

ARTICLE INFO

Article history:
Received 28 October 2010
Received in revised form
29 August 2011
Accepted 22 October 2011
Available online 28 March 2012

Keywords:
Networked control system (NCS)
Q-function
Adaptive estimator
Optimal control

ABSTRACT

In this paper, the stochastic optimal control of linear networked control system (NCS) with uncertain system dynamics and in the presence of network imperfections such as random delays and packet losses is derived. The proposed stochastic optimal control method uses an adaptive estimator (AE) and ideas from Q-learning to solve the infinite horizon optimal regulation of unknown NCS with time-varying system matrices. Next, a stochastic suboptimal control scheme which uses AE and Q-learning is introduced for the regulation of unknown linear time-invariant NCS that is derived using certainty equivalence property. Update laws for online tuning the unknown parameters of the AE to obtain the Q-function are derived. Lyapunov theory is used to show that all signals are asymptotically stable (AS) and that the estimated control signals converge to optimal or suboptimal control inputs. Simulation results are included to show the effectiveness of the proposed schemes. The result is an optimal control scheme that operates forward-in-time manner for unknown linear systems in contrast with standard Riccati equation-based schemes which function backward-in-time.

1. Introduction

Feedback control systems with control loops closed through a real-time network are called Networked Control Systems (NCSs) (Antsaklis & Baillieul, 2004; Branicky, Phillips, & Zhang, 2000; Cloosterman, van de Wouw, Heemels, & Nijmeijer, 2009; Halevi & Ray, 1988; Wu & Chen, 2007). In NCS, a communication packet carries the reference input, plant output, and control input which are exchanged using a network among control system components such as sensors, controller, and actuators. The primary advantages of NCS are reduced system wiring, ease of system diagnosis and maintenance, and increased system agility. However, insertion of the communication network in the feedback loop brings many challenging issues.

The first issue is the network-induced delay that occurs while exchanging data among devices connected to the shared medium. This delay, either constant or random, can degrade the performance of control system and even destabilize the system when the delay is not explicitly considered in the design process. The second issue is packet losses due to unreliable network transmission which can cause a loss in control input resulting in instability. These issues have been identified in the literature and are being studied.

For instance, Cloosterman et al. (2009) analyzed the stability of NCS with network-induced delays. Walsh, Ye, and Bushnell (1999) and Lian, Moyne, and Tilbury (2003) considered stability performance of NCS with constant delays. Azimi-Sadjadi (2003), Wu and Chen (2007) and Schenato, Sinopoli, Franceschetti, Poolla, and Sastry (2007) analyzed the stability performance of NCS with packet losses. Eventually Zhang, Branicky, and Phillips (2001) conducted the stability analysis of NCS with delays and packet losses and proposed a stability region.

While stable controllers are encouraging, optimality is generally preferred for NCS which is very difficult to attain. Lian et al. (2003) proposed the optimal controller design by using classical optimal control theory (Lewis & Syrmos, 1995) for NCS with multiple constant delays embedded into the NCS representation. Using the stochastic optimal control theory (Åstrom, 1970; Bertsekas & Shreve, 1978; Stengel, 1986), Nilsson, Bernhardsson, and Wittenmark (1998) proposed the optimal and suboptimal controller designs for linear NCS with random delays. Although these optimal and suboptimal controller designs have resulted in satisfactory performance, they all require information about the system

☆ Research in part by NSF grant ECCS# 1128281 and Intelligent Systems Center. The material in this paper was partially presented at the 2011 American Control Conference, June 29–July 1, 2011, San Francisco, California, USA. This paper was recommended for publication in revised form by Associate Editor Gang Tao under the direction of Editor Miroslav Krstic.

E-mail addresses: hx6h7@mst.edu (H. Xu), sarangap@mst.edu (S. Jagannathan), lewis@uta.edu (F.L. Lewis).

[1] Tel.: +1 573 578 2572; fax: +1 573 578 2572.

0005-1098/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2012.03.007

NCS dynamics and information on delays and packet losses which are not known beforehand.

On the other hand, adaptive dynamic programming (ADP) schemes proposed by Werbos (1991), Barto, Sutton, and Anderson (1983), Busoniu, Babuska, and Schutter (2010), and Watkins (1989), intend to solve optimal control problems forward-in-time by using value and policy iterations. There are four techniques in ADP (i.e. heuristic dynamic programming (HDP), action dependent HDP (ADHDP), dual heuristic programming (DHP) and action dependent DHP (ADDHP)), but they all require policy and value iterations.

Al-Tamimi, Lewis, and Abu-Khalaf (2007) used the Q-learning policy iteration method to solve the optimal strategies for linear discrete-time system quadratic zero-sum games in forward-in-time without requiring the system dynamics wherein the system dynamics are defined as constant matrices. It is important to note that policy and value iteration-based schemes are difficult to implement on hardware (Dierks, Thumati, & Jagannathan, 2009) since it is not clear how to select the number of iterations required for convergence and stability while keeping the hardware constraints. Inadequate number of policy and value iterations can result in instability (Dierks et al., 2009). Therefore, Dierks and Jagannathan (2009) used two time-based neural networks (NN) to solve the Hamilton–Jacobi–Bellman (HJB) equation forward-in-time for the optimal control of a class of general nonlinear affine discrete-time systems without using policy or value iterations. However, these papers did not consider the effects of delays and packet losses which are normally found in a NCS. The delays and packet losses cause instability (Zhang et al., 2001) if they are not considered carefully which in turn make the optimal controller design more involved and different than Al-Tamimi et al. (2007) and Zhang, Luo, and Liu (2009).

Thus, this paper introduces ADHDP technique for the optimal and suboptimal control of linear NCS with uncertain system dynamics and in the presence of unknown random network-induced delays and packet losses. In other words, first a linear NCS with random delays and packet losses will be represented by a time-varying linear system with unknown system matrices. The suboptimal approach in Al-Tamimi et al. (2007); Zhang et al. (2009) is not directly applicable to the NCS due to the inclusion of network imperfections such as these delays and packet losses.

A novel approach is undertaken to the optimal regulation of linear NCS with random delays and packet losses to solve the Bellman equation (Wonham, 1968) online and forward-in-time without using policy and value iterations. Using an initial stabilizing control, an adaptive estimator (AE) (Franklin, Powell, & Emani-Naeini, 1994) is tuned online to learn the stochastic cost function without needing to solve the stochastic Riccati equation (SRE). Then, using the idea of Q-learning, the optimal controller which minimizes the stochastic cost function can be calculated based on the information provided by the AE. Thus the proposed AE-based scheme relaxes the requirements for system dynamics and information on random delay and packet losses. Next, the suboptimal controller design is derived based on NCS representation that is obtained by using certainty equivalence property. For the suboptimal control, linear NCS is modeled as a time-invariant system with unknown matrices. The suboptimal controller reduces computational complexity.

This paper is organized as follows. First, NCS background representation is given in Section 2. In Section 3, the stochastic optimal and suboptimal regulation controls of NCS are introduced. Section 4 illustrates the effectiveness of proposed schemes via numerical simulations, and Section 5 provides concluding remarks.

2. Background

The basic structure of NCS considered in this paper is shown as Fig. 1 where the feedback control loop is closed over a wireless network. Since wireless network bandwidth is limited, two types of network-induced delays and one type of packet losses are included in this structure: (1) $\tau_{sc}(t)$: sensor-to-controller delay, (2) $\tau_{ca}(t)$: controller-to-actuator delay, and (3) $\gamma(t)$: indicator of packet received.

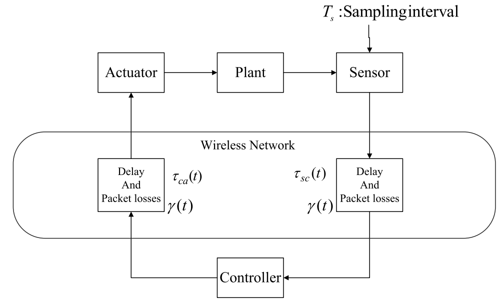

Fig. 1. Networked control system (NCS).

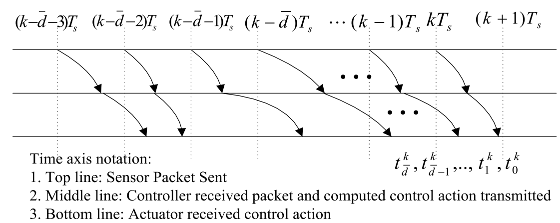

Fig. 2. Timing diagram of signals in NCS.

The following assumption is needed similar to other works (Hu & Zhu, 2003; Liou & Ray, 1991):

Assumption 1. (a) Sensor is time-driven while the controller and actuator are event-driven (Hu & Zhu, 2003).

(b) The communication network considered is a wide area wireless network so that the two types of network-induced delays are independent and unknown whereas their probability distribution functions are considered known (Goldsmith, 2003; Hu & Zhu, 2003; Liou & Ray, 1991).

(c) The sum of the two delay types is bounded (Liou & Ray, 1991) while the initial state of linear system is deterministic (Hu & Zhu, 2003).

A linear time-invariant system $\dot{x}(t) = Ax(t) + Bu(t)$ is considered. However, considering the effects of network-induced delays and packet losses, the original controlled plant can be expressed as

$$\dot{x}(t) = Ax(t) + \gamma(t)Bu(t - \tau(t)) \tag{1}$$

where

$$\gamma(t) = \begin{cases} \mathbf{I}^{n\times n} & \text{if the control input is received at time } t \\ \mathbf{0}^{n\times n} & \text{if the control input is lost at time } t, \end{cases}$$

$x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ representing system matrices. From Assumption 1, we can assume that the sum of network-induced delays is bounded above i.e. $\tau(t) = \tau_{sc}(t) + \tau_{ca}(t) < \bar{d}T_s$ where $\bar{d}$ represents the delay bound while T_s is the sampling interval.

During a sampling interval $[kT_s, (k+1)T_s)$ $\forall k$, the controller input $u(t)$ to the plant is a piecewise constant. According to Assumption 1, there are at most $\bar{d}$ current and previous control input values that can be received at the actuator. If several control inputs are received at the same time, only the newest control input is allowed to act on the controlled plant during any sampling interval $[kT_s, (k+1)T_s)$ $\forall k$, and other previous control inputs are deduced. Since controller is event driven, the plant will implement control input at these time instant $kT_s + t^k_i$, $i = 0, 1, \ldots, \bar{d}$ and $t^k_i < t^k_{i-1}$ where $t^k_i = \tau^k_i - iT_s$ as illustrated in Fig. 2 (Liou & Ray, 1991).