Contents lists available at ScienceDirect

Automatica

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

Brief paper

Model-free event-triggered control algorithm for continuous-time linear systems with optimal performance☆

Kyriakos G. Vamvoudakis [a], Henrique Ferraz [b]

[a] Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA, USA
[b] Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA, USA

ARTICLE INFO

Article history:
Received 10 November 2015
Received in revised form
21 February 2017
Accepted 21 February 2017
Available online xxxx

Keywords:
Networked control
Optimal control
Q-learning

ABSTRACT

This paper proposes a new model-free event-triggered optimal control algorithm for continuous-time linear systems. The problem is formulated as an infinite-horizon optimal adaptive learning problem, and we are able to simultaneously address the issue of designing a control and a triggering mechanism with guaranteed optimal performance by design. In order to provide a model-free solution, we adopt a Q-learning framework with a critic network to approximate the optimal cost and a zero-order hold actor network to approximate the optimal control. Since we have dynamics that evolve in continuous and discrete-time, we write the closed-loop system as an impulsive model and prove asymptotic stability of its equilibrium. Numerical simulation of an unknown unstable system is presented to show the efficacy of the proposed approach.

1. Introduction

In conventional implementations of control systems, control tasks are usually executed by periodically sampling the plant's output and updating the control inputs. Selection of sampling period is traditionally performed during the control design stage, having in mind the trade off between the reconstruction of the continuous-time signal and the load on the computer (Astrom & Willenmark, 1997). In many applications, however, conventional design methods may not represent an efficient solution. In network control systems where communication channels are shared with multiple and possibly remotely located sensors, actuators and controllers (Hespanha, Naghshtabrizi, & Xu, 2007), periodic and high frequency sampling, computation and transmission of data could result in inefficient use of resources of bandwidth and energy.

In systems with limited bandwidth, event-triggered control (Astrom & Bernhardsson, 1999; Heemels, Johansson, & Tabuada, 2012; Tabuada, 2007) and self-triggered control (Anta & Tabuada, 2010; Gommans, Antunes, Donkers, Tabuada, & Heemels, 2014; Wang & Lemmon, 2009) represent two emerging control strategies that have been shown to be suitable for reducing the communication between actuators/sensors, and the controller. This is attained by letting the system evolve in open-loop and only closing the loop whenever a user designed triggering condition that guarantees stability and performance is satisfied. Sparse communication and less computation could result in decongestion of the network channels and energy save for devices.

While the event-triggered control and the self-triggered control literature continues to flourish, two fundamental issues remain overshadowed, as pointed out in Gommans et al. (2014): the co-design of both the feedback law and the triggering scheme; and the performance guarantees by design for the proposed algorithm. So far, to the best knowledge of the authors, only a few approaches have tried to simultaneously address these points. For example, in Gommans et al. (2014), an optimal self-triggered control with discounted quadratic cost function for discrete time linear systems is proposed. It is shown that in some cases the sparse communication strategy could outperform the traditional periodic time-triggered one. In Peng and Yang (2013), the authors explore another approach to the problem of the co-design with an H_∞ performance index for network control systems with communication delays and packet losses. However, these methods rely on an offline computation of the Riccati or Hamilton–Jacobi–Bellman equation and depend on the *full knowledge of the system dynamics, being vulnerable to exhaustive modeling and malicious attacks.*

☆ The work was partially supported by a Virginia Tech startup fund and by a CAPES BEX 1111-13-1 grant. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Akira Kojima under the direction of Editor Ian R. Petersen.

E-mail addresses: kyriakos@vt.edu (K.G. Vamvoudakis), henrique@ece.ucsb.edu (H. Ferraz).

http://dx.doi.org/10.1016/j.automatica.2017.03.013

The combination of optimal control theory (Lewis & Syrmos, 1995) and adaptive control theory (Ioannou & Fidan, 2006) can be brought together with the use of ideas from reinforcement learning (Sutton & Barto, 1998) to overcome these issues. Approximate dynamic programming has been shown to be a powerful tool to solve reinforcement learning problems in an adaptive way and also to guarantee optimal performance, (Busoniu, Babuska, deSchutter, & Ernst, 2010; Powell, 2007; Vrabie, Vamvoudakis, & Lewis, 2012; Zhang, Liu, Luo, & Wang, 2012). Actor/critic algorithms are a form of reinforcement learning (Sutton & Barto, 1998) which uses an actor structure to select the control policies to improve the performance and a critic structure to evaluate actor's decisions.

Recently, some studies have tried to combine event-triggered control algorithms for problems where the system dynamics is unknown. One of the earliest attempts was presented in Arzen (1999) and further developed in Durand and Marchand (2009) and in Wang, Mounier, Cela, and Niculescu (2011) where the authors used a PID type of controller with event-based updates. Despite its inherently simple structure and easiness to tune, these types of controllers do not provide any optimality guarantees. As of event-triggered control algorithms for unknown systems with optimality guarantees, Sahoo, Xu, and Jagannathan (2017) presents an algorithm where a neural network based identifier is used to approximate the unknown nonlinear continuous-time system. The resulting closed-loop signals are locally ultimately bounded and the controller is near-optimal. For the cases where the full state information is not available, Zhong and He (2017) proposes a scheme combining neural network based observer with an optimal event-triggered control algorithm. As a result of the approximation, the stability result obtained is local and the closed-loop signals are also ultimately bounded.

In our previous work, in Vamvoudakis (2014), we have derived a novel optimal adaptive event triggered control algorithm for known nonlinear systems by using an approach based on Hamiltonians. This did not enable us to define a model-free approach. For that reason, in this paper, we derive a novel model-free approach based on Q-learning while also guaranteeing that the Zeno behavior is excluded. Q-learning is a model-free reinforcement learning (Bertsekas & Tsitsiklis, 1996; Busoniu et al., 2010; Powell, 2007; Sutton & Barto, 1998; Zhang et al., 2012) technique primarily developed for discrete-time systems where an optimal action is selected based on previous state and actions observations (Watkins & Dayan, 1992). It learns an action-dependent value function that ultimately gives the expected utility of taking a given action in a given state and following the optimal policy thereafter. When such an action-dependent value function is learned, then the optimal policy can be computed easily. The biggest strength of Q-learning is that it does not require a model of the system to be controlled.

Main results: The contributions of the paper are threefold. We first show that the optimal event-triggered control policy is suboptimal with respect to the time-triggered optimal control one. Further, in order to derive a scheme that is independent of the system matrices, we use Q-learning and derive appropriate tuning laws to learn the newly proposed Q-function. Specifically, we use an actor/critic structure that adaptively tunes and approximates the optimal event-triggered controller and the Q-function, respectively, to solve the problem online and forward in time. Finally, by using an impulsive systems model for the closed-loop system, we prove that the equilibrium point of the flow and the jump dynamics is globally asymptotically stable.

Structure: This paper is structured as follows. Section 2 formulates the infinite horizon optimal control problem, Section 3 provides a brief background on the optimal control solution and the relationship between the optimal time and even-triggered control policies. Since Section 3 relies on complete knowledge of the system matrices, Section 4 provides a model-free formulation based on a Q-learning approach and an actor/critic structure to estimate the parameters of the Q-function. Rigorous Lyapunov based proof of asymptotic stability is provided and the existence of a positive lower bound for the inter-event times is shown. Numerical simulations are presented in Section 5 to show the efficacy of the proposed algorithm and finally Section 6 concludes and talks about future work.

Notation: The notation used here is standard. $\mathbb{R}^+$ is the set of positive real numbers. We denote $\underline{\lambda}(A)$ and $\bar{\lambda}(A)$ as the minimum and maximum, respectively, eigenvalues of a matrix A. Also, $\|\cdot\|$ denotes the Euclidean norm for a vector and the Frobenius norm for a matrix. The Kronecker product is represented by $\otimes$ and the half-vectorization, vech(A), of a symmetric $n \times n$ matrix A is the $n(n+1)/2$ column vector obtained by vectorizing only the lower (or upper) triangular part of A, vech$(A) = [A_{1,1} \cdots A_{n,1}\ A_{2,2} \cdots A_{n,2} \cdots A_{n-1,n-1}\ A_{n-1,n}\ A_{n,n}]^T$ and mat$(\cdot)$ is the inverse vech operation known as matrization. The colon symbol: can be used to form implicit vectors from a matrix or vector, i.e. $A(j:k), k > j$ is $[A(j)\ A(j+1) \cdots A(k)]$. A function $\kappa : \mathbb{R}^+ \to \mathbb{R}$ is said to belong to class $\mathcal{K}$ functions if it is continuous, strictly increasing and $\kappa(0) = 0$. Through out this work, we will use the closed-loop impulsive system formulation as in Haddad, Chellaboina, and Nersesov (2006) and Hespanha, Liberzon, and Teel (2008) defined as follows:

$$\begin{cases} \dot{\psi} = f(\psi), & \forall t \in (r_j, r_{j+1}] \\ \psi^+ = g(\psi), & t = r_j, \end{cases}$$

where $\{r_j\}_{j=0}^{\infty}$ is a monotonically increasing sequence of sampling instants with r_j the jth consecutive sampling instant satisfying $\lim_{j\to\infty} r_j = \infty$; the state $\psi \in \mathbb{R}^{n_\psi}$ is continuous between the sampling instants; f and g are the flow and the jump dynamics, respectively, and from $\mathbb{R}^{n_\psi}$ to $\mathbb{R}^{n_\psi}$. We denote by $(\cdot)^+$ the right-limit operator, i.e., $\psi^+ = \lim_{s\searrow t} \psi(s)$.

2. Problem formulation

Consider the following linear time invariant continuous-time system,

$$\dot{x}(t) = Ax(t) + Bu(t), \quad x(0) = x_0,\ t \geqslant 0 \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is a measurable state vector, $u(t) \in \mathbb{R}^m$ is the control input and $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ are the plant and input matrices, respectively, that will be considered uncertain/unknown.

To save resources, the controller will work with a sampled version of the state defined as follows:

$$\hat{x}(t) = \begin{cases} x(r_j), & \forall t \in (r_j, r_{j+1}] \\ x(t), & t = r_j. \end{cases}$$

The controller maps the sampled state onto a control vector which after using a zero-order hold becomes a continuous-time input signal. In order to decide when to trigger an event, we will define the gap between the current state $x(t)$ and the sampled state $\hat{x}(t)$ as

$$e(t) := \hat{x}(t) - x(t). \tag{2}$$

Remark 1. Note that when an event is triggered at $t = r_j, j \in \mathbb{N}$, a new state measurement is rendered that resets the gap (2) to zero. □

In this work, we are interested in finding a controller $u(t) = u_d$ of the form $u_d := k(\hat{x}(t))$ that minimizes the cost functional, $J(x(0); u_d) = \frac{1}{2}\int_0^\infty \left(x^T Hx + u_d^T Ru_d\right) d\tau$, with user-defined matrices $H \succeq 0$, $R \succ 0$, and with reduced updates of the control input given by the triggering rule.