Information Sciences 317 (2015) 96–113

Contents lists available at ScienceDirect

Information Sciences

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

Optimal distributed synchronization control for continuous-time heterogeneous multi-agent differential graphical games

Qinglai Wei [a], Derong Liu [a,*], Frank L. Lewis [b,c]

[a] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] UTA Research Institute, University of Texas at Arlington, Fort Worth, TX, USA
[c] State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, China

A R T I C L E   I N F O

Article history:
Received 4 November 2014
Received in revised form 9 March 2015
Accepted 18 April 2015
Available online 24 April 2015

Keywords:
Adaptive critic designs
Adaptive dynamic programming
Approximate dynamic programming
Heterogeneous multi-agents
Graphical games
Policy iteration

A B S T R A C T

In this paper, a new optimal distributed synchronization control scheme for the consensus problem of heterogeneous multi-agent differential graphical games is developed by iterative adaptive dynamic programming (ADP). The main idea is to use iterative ADP technique to obtain the iterative control law which makes all the agents track a given dynamics and simultaneously makes the iterative value function reach the Nash equilibrium. In the developed heterogeneous multi-agent differential graphical games, the agent of each node is different from one another. The dynamics and performance index function for each node depend only on local neighborhood information. A cooperative policy iteration algorithm is presented to achieve the optimal distributed synchronization control law for the agent of each node, where the coupled Hamilton–Jacobi equations for optimal synchronization control of heterogeneous multi-agent differential games can be avoided. Convergence analysis is developed to show that the iterative value functions of heterogeneous multi-agent differential graphical games can converge to the Nash equilibrium. Two simulation examples are given to show the effectiveness of the developed optimal control scheme.

1. Introduction

Intelligent control has always been the key focus in the control field in the latest few decades [30,38,41,42,48]. A large class of control systems are controlled by more than one controller or decision maker with each using an individual strategy. These controllers often operate in a group with coupled performance index functions as a game [17]. Stimulated by a vast number of applications, including those in economics, management, communication networks, power networks, and complex engineering systems, game theory [36] has been very successful in modeling strategic behaviors, where the outcome of each player depends on the actions of himself and all the other players. In [9], considering the disturbance as a control input of the system, the optimal control problem with disturbance was transformed into a two-person zero-sum games and the existence criterion of the optimal solution (which is the saddle point) of zero-sum games was also proposed. Hence, zero-sum games have been widely used in the optimal control for systems with disturbances [2,3,8,21,60,64]. In many real world control systems, there exist multi-controllers in the systems and the performance index functions for these controllers are not zero-sum, such as the multi-agent graphical games. Synchronization behavior, found in flocking of birds, schooling of

* Corresponding author.
E-mail addresses: qinglai.wei@ia.ac.cn (Q. Wei), derong.liu@ia.ac.cn (D. Liu), lewis@uta.edu (F.L. Lewis).

http://dx.doi.org/10.1016/j.ins.2015.04.044

fish, and other natural systems, is an primary behavior in multi-agent systems. Distributed synchronization control, which drives all the agent to a desired synchronization behavior, becomes an important research topic for the control of the multi-agent systems [15,30,43,44,49]. However, in the previous literature on distributed synchronization control problems for multi-agent systems, the stability of the multi-agent systems are focused, while the optimality property, which is also an primary property for multi-agent systems, are scarce to investigate. The main difficulty lies in solving the coupled Hamilton–Jacobi (HJ) equations. To overcome these difficulties, new effective optimal control schemes need to developed.

Adaptive dynamic programming (ADP) [58,59], characterized by strong abilities of self-learning and adaptivity, has received significantly increased attention and becomes an important brain-like intelligent optimal control method [12,14,29,35,57,65–67]. There were several synonyms used for ADP, including "adaptive critic designs" [22,37], "adaptive dynamic programming" [25,52], "approximate dynamic programming" [33], "neuro-dynamic programming" [10], "neural dynamic programming" [11], and "reinforcement learning" [13,16,39,61,62]. Iterative methods, including value and policy iterations [20], are widely used in ADP to obtain the optimal performance index function indirectly [7,18,23–25,27,51,53–56]. Value iteration for optimal control was first given in [10]. In [6], the convergence of value iteration ADP method was proven. In [4,5], value iteration algorithms were proposed to solve two-person zero-sum games for discrete-time linear systems. For value iteration algorithms, the stability of system under the iterative control law cannot be guaranteed. Policy iteration algorithms are a class of the most primary and important iterative ADP algorithms [19,31,32,50]. Policy iteration for optimal control of continuous-time systems was given in [1,34]. In [40], policy iteration algorithm was successfully applied to solve continuous-time complex-valued systems. Discrete-time policy iteration with convergence and stability proofs was developed in [28]. In [66], a continuous-time two-person zero-sum game with general utility function was solved by policy iteration. In [21], integral reinforcement learning based policy iteration algorithm was proposed to obtain optimal control laws for linear continuous-time zero-sum games with completely unknown dynamics. In [26], policy iteration was successfully applied to achieve optimal control laws for multi-person zero-sum differential games. Policy iteration algorithm for non-zero-sum games was proposed in [45,46]. In [63], a near-optimal control for nonzero-sum differential games was solved using single-network policy iteration algorithm. In previous policy iteration ADP algorithms for multi-player games, it is always desired that the system states for each agent converge to the equilibrium of the systems. In many real world games, it requires that the states of each agent track a desired dynamics, i.e., to achieve synchronization control. Synchronization behavior of the multi-agent optimal control based on ADP was pioneered by Vamvoudakis et al. [47]. In [47], policy iteration algorithm was developed for multi-agent graphical differential games, where the Nash equilibrium of the game was effectively achieved. However, in [47], to guarantee the convergence of performance index function of the agent for each node, it requires that all the dynamics of autonomous systems for each node of the multi-agent graphical games are the same, which equal the desired dynamics. If the autonomous systems of each agent are not the same, i.e., for heterogeneous multi-agent graphical differential games, the convergence analysis in [47] is invalid. This is a disadvantage of the proposed policy iteration algorithm in [47]. To the best of the authors' knowledge, when the dynamics of the graphical game are different, the optimal synchronization control for the synchronization of heterogeneous multi-agent graphical differential games is not discussed. This motivates our research.

In this paper, inspired by the pioneering work of [47], an optimal distributed synchronization control for heterogeneous multi-agent graphical differential games is investigated. We emphasize that in the developed heterogeneous graphical differential games, the autonomous system of the agent for each node and the desired dynamics can be different from each other. Using system transformations, the optimal synchronization control problem is transformed into an optimal multi-agent regulation problem. The main contribution of this paper is to develop an effective cooperative policy iteration algorithm to obtain optimal distributed synchronization control law for heterogeneous multi-agent graphical differential games. Convergence properties of the cooperative policy iteration algorithm are developed, which guarantee that the iterative value function of the agent for each node converges to the Nash equilibrium of the games. Two simulation examples are presented to show the effectiveness of the developed algorithm.

This paper is organized as follows. In Section 2, graphs and synchronization of heterogeneous multi-agent dynamic systems are presented. In Section 3, the optimal distributed cooperative control for heterogeneous multi-agent differential graphical games is presented. The optimal control law is developed in this section. In Section 4, heterogeneous multi-agent differential graphical games by cooperative policy iteration algorithm are developed. Properties of the iterative value functions are analyzed. In Section 5, simulation results are given to demonstrate the effectiveness of the developed algorithm. Finally, in Section 6, conclusions will be given and our future work is declared.

2. Graphs and synchronization of multi-agent systems

In this section, a background review of communication graphs is given and the problem of synchronization of heterogeneous multi-agent systems is formulated.

2.1. Graph theory

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ be a weighted graph of N nodes with the nonempty finite set of nodes $\mathcal{V} = \{v_1, \ldots, v_N\}$, where the set of edges $\mathcal{E}$ belongs to the product space of $\mathcal{V}$ (i.e., $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$), an edge of $\mathcal{G}$ is denoted by $\rho_{ij} = (v_j, v_i)$, which is a direct path