SCIENCE CHINA
Information Sciences

• LETTER •

October 2020, Vol. 63 209205:1–209205:3
https://doi.org/10.1007/s11432-019-9868-2

Optimal containment control of continuous-time multi-agent systems with unknown disturbances using data-driven approach

Zhinan PENG[1], Jiefu ZHANG[1], Jiangping HU[1*],
Rui HUANG[1] & Bijoy Kumar GHOSH[1,2]

[1]*School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu* 611731*, China;*
[2]*Department of Mathematics and Statistics, Texas Tech University, Lubbock* 79409*, USA*

Received 27 February 2019/Accepted 28 March 2019/Published online 14 April 2020

Citation Peng Z N, Zhang J F, Hu J P, et al. Optimal containment control of continuous-time multi-agent systems with unknown disturbances using data-driven approach. Sci China Inf Sci, 2020, 63(10): 209205, https://doi.org/10.1007/s11432-019-9868-2

Dear editor,
Containment control problems of multi-agent systems (MASs) have attracted considerable attention in recent years owing to their widespread use in applications pertaining to spacecraft control, sensor networks, power systems, and mobile robots.

The containment control problem was solved for MASs with high-order dynamics and random switching networks respectively in [1, 2]. The containment control was designed for multi-rigid systems in [3]. In [4], the containment control was discussed for heterogeneous linear MASs. However, most existing containment control methods rely on explicit system models, which are often quite difficult to obtain in the real world. It has been shown that this problem may be efficiently resolved by using data-driven reinforcement learning technology [5]. In [6], a data-driven value iteration algorithm was proposed to solve a containment control problem for discrete-time MASs. The optimal containment control problem of continuous-time MASs with disturbances remains to be investigated. In fact, the disturbances in the system, which are a common facet of many real applications, pose a challenge.

Motivated by the above observations, this study addresses an optimal containment control problem for continuous-time MASs with unknown disturbances. The main contributions are summarized as follows. (1) An appropriate model transformation is presented to convert the original containment control problem of MASs with disturbances into an optimal containment control problem, where the unknown disturbances no longer exist. Thus, the data-driven learning-based approach comes into play. (2) The equivalence relation between the original MASs with unknown disturbances and the converted optimal containment control system is presented. (3) A data-driven based neural network framework is adopted during the implementation of the proposed approach, negating the need for an accurate system dynamic model.

Problem formulation. Consider an MAS consisting of n followers and m leaders, which are denoted by $\{1, 2, \ldots, n\}$ and $\{n+1, n+2, \ldots, n+m\}$, respectively. The followers' dynamics are described by

$$\dot{x}_i(t) = Ax_i(t) + B_i u_i(t) + D_i \omega_i(t), \qquad (1)$$

where $x_i(t) \in \mathbb{R}^N$ is the state of follower i and $u_i(t) \in \mathbb{R}^{q_i}$ is the control input. $\omega_i \in \mathbb{R}^{l_i}$ presents unknown disturbances bounded by a non-negative

* Corresponding author (email: hujp@uestc.edu.cn)

known function τ_b. Notice that ω_i represents unmatched disturbances, that is, $B_i \neq D_i$ for $\forall i$. The leaders' dynamics are described by

$$\dot{x}_0^{(k)}(t) = Ax_0^{(k)}(t), \quad k = n+1, \ldots, n+m, \quad (2)$$

where $x_0^{(k)} \in \mathbb{R}^N$ is the state of the kth leader. Herein we assume the system matrix A is unknown.

The objective of containment control is to identify control policies to ensure that all the followers can converge to the convex hull spanned by the leaders as $t \to \infty$, that is,

$$\lim_{t\to\infty} \mathrm{dist}\big(x_i(t), \mathrm{Co}(x_0^{(k)})\big) = 0, \quad (3)$$

where $\mathrm{Co}(x_0^{(k)})$ denotes the convex hull of a finite point set $\{x_0^{(k)} | k = n+1, \ldots, n+m\}$. $\mathrm{dist}(x, \mathcal{C})$ denotes the distance from $x \in \mathbb{R}^N$ to the set $\mathcal{C} \subseteq \mathbb{R}^N$, that is, $\mathrm{dist}(x, \mathcal{C}) = \inf_{y\in\mathcal{C}} \|x - y\|$.

The interaction relationships among followers can be described by a weighted adjacency matrix $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{n\times n}$. The connection weight between the kth leader and the ith follower is denoted by g_i^k. Then, $g_i^k > 0$ if the ith follower is connected to the kth leader; otherwise, $g_i^k = 0$. Then, the adjacency matrix of the kth leader is $\mathcal{B}_k = \mathrm{diag}\{g_1^k, \ldots, g_n^k\}$. Let N_i be the neighbors of agent i. The Laplacian matrix is given by $\mathcal{L} = \mathcal{D} - \mathcal{A}$, where $\mathcal{D} = \mathrm{diag}\{d_i\}$ is a diagonal matrix, $d_i = \sum_{j\in N_i} a_{ij}$.

Define the local neighbor tracking error of the ith follower as

$$e_i(t) = \sum_{j\in N_i} a_{ij}(x_i - x_j) + \sum_{k=n+1}^{n+m} g_i^k (x_i - x_0^{(k)}). \quad (4)$$

We can further rewrite (4) in the compact form, $e(t) = \sum_{k=n+1}^{n+m} (\mathcal{H}_k \otimes I_N)(x - \bar{x}_k)$, where $\bar{x}_k = 1_n \otimes x_0^{(k)}$ and $\mathcal{H}_k = \frac{1}{m}\mathcal{L} + \mathcal{B}_k$.

Consider the dynamics of the local neighborhood tracking error for agent i as

$$\dot{e}_i = Ae_i + \left(d_i + \sum_{k=n+1}^{n+m} g_i^k\right) B_i u_i + \left(d_i + \sum_{k=n+1}^{n+m} g_i^k\right) \cdot D_i\omega_i - \sum_{j\in N_i} a_{ij} B_j u_j - \sum_{j\in N_i} a_{ij} D_j \omega_j. \quad (5)$$

According to [4], $\|e(t)\| \to 0$ implies the convergence of followers to the convex hull spanned by the leaders as $t \to \infty$, that is, the containment control is achieved.

Implementation of the optimal containment control policy design. In order to establish the transformed optimal control system for any i, we divide the uncertainty term $D_i\omega_i$ into two parts: a matched component and an unmatched component. By projecting $D_i\omega_i$ onto the range of B_i, we have

$$D_i\omega_i = B_i B_i^{\dagger} D_i \omega_i + (I - B_i B_i^{\dagger}) D_i \omega_i, \quad (6)$$

where $B_i^{\dagger}$ is the Moore-Penrose pseudoinverse of B_i.

Therefore, a converted MAS can be constructed as follows:

$$\dot{x}_i = Ax_i + B_i u_i + (I - B_i B_i^{\dagger}) D_i v_i, \quad (7)$$

where the vector $[u_i; v_i]$ is the control policy of the converted MAS (7), and $v_i \in \mathbb{R}^{q_i}$ is an augmented control policy and used to handle the unmatched uncertainty component.

According to (4) and (7), the dynamics (5) can be rewritten as

$$\dot{e}_i = Ae_i + \left(d_i + \sum_{k=n+1}^{n+m} g_i^k\right) (B_i u_i + (I - B_i B_i^{\dagger}) D_i v_i) - \sum_{j\in N_i} a_{ij} (B_j u_j + (I - B_j B_j^{\dagger}) D_j v_j). \quad (8)$$

The neighborhood tracking error dynamics contain multiple control signals, that is, u_i, v_i of agent i and their neighbors.

Define the local performance indices of (7) as

$$V_i = \int_0^{\infty} (\zeta^2 \|p_{ii}\|^2 \tau_b^2 + \|\rho\|^2 + r(e_i(\tau), u_i(s), u_{-i}(s), v_i(s), v_{-i}(s)))\mathrm{d}s, \quad (9)$$

where $r(e_i, u_i, u_{-i}, v_i, v_{-i})$ is the utility function and defined by

$$\begin{aligned} &r(e_i, u_i, u_{-i}, v_i, v_{-i}) \\ &= e_i^{\mathrm{T}} Q_{ii} e_i + u_i^{\mathrm{T}} R_{ii} u_i + \zeta^2 v_i^{\mathrm{T}} P_{ii} v_i \\ &\quad + \sum_{j\in N_i} u_j^{\mathrm{T}} R_{ij} u_j + \sum_{j\in N_i} \zeta^2 v_j^{\mathrm{T}} P_{ij} v_j, \end{aligned}$$

the matrices $Q_{ii} > 0$, $R_{ii} > 0$, $P_{ii} > 0$, $R_{ij} > 0$, and $P_{ij} > 0$ are positive definite, $P_{ii} = p_{ii} p_{ii}^{\mathrm{T}}$, p_{ii} is an invertible matrix, ρ and $\zeta > 0$ are the designed parameters, and $u_{-i} = \{u_j | j \in N_i\}$ and $v_{-i} = \{v_j | j \in N_i\}$ are respectively the sets of control policies from the neighbors of agent i.

Then the Hamiltonian function H_i can be defined as

$$\begin{aligned} H_i = &(\nabla V_i)^{\mathrm{T}} \dot{e}_i + \|\rho\|^2 + \zeta^2 \|p_{ii}\|^2 \tau_b^2 \\ &+ e_i^{\mathrm{T}} Q_{ii} e_i + u_i^{\mathrm{T}} R_{ii} u_i + \zeta^2 v_i^{\mathrm{T}} P_{ii} v_i \\ &+ \sum_{j\in N_i} u_j^{\mathrm{T}} R_{ij} u_j + \sum_{j\in N_i} \zeta^2 v_j^{\mathrm{T}} P_{ij} v_j = 0, \quad (10) \end{aligned}$$

where the boundary condition is $V_i = 0$, and $\nabla V_i = \partial V_i / \partial e_i$ is the partial derivative of the performance index V_i with respect to e_i.

The optimal control policies u_i^* and v_i^* satisfy the first-order necessary condition and are described as

$$u_i^* = -\frac{1}{2}\left(d_i + \sum_{k=n+1}^{n+m} g_i^k\right) R_{ii}^{-1} B_i^{\mathrm{T}} \nabla V_i^*, \quad (11)$$

$$v_i^* = -\frac{1}{2\zeta^2}\left(d_i + \sum_{k=n+1}^{n+m} g_i^k\right) P_{ii}^{-1} D_i^{\mathrm{T}} \cdot (I - B_i B_i^{\dagger}) \nabla V_i^*, \quad (12)$$

where V_i^* denotes the optimal performance indices.

The following theorem gives the equivalence relation of the original system with disturbances and the converted system under the proposed method. The proof of the theorem is given in Appendix A.

Theorem 1. Consider system (7) with performance index (9) for agent i. The optimal control policies u_i^* in (11) are also the solutions of the original containment control problem, if the following conditions are satisfied, that is, $e_i^{\mathrm{T}} Q_{ii} e_i > v_i^{*\mathrm{T}} P_{ii} v_i^*$, and $0 < \zeta < \sqrt{2}/2$.

Data-driven based neural network framework. The above optimal containment control policy design is implemented via a data-driven based critic-actor network, which is established as follows. The critic-network of agent i uses tracking error e_i, control input u_i, and its neighbors' control input u_{-i} as the input for this network. The output is an approximation of performance indices $\hat{V}_i$, that is,

$$\hat{V}_i(e_i) = \omega_{ci2}^{\mathrm{T}} \Phi(z_{ci}(e_i)), \quad (13)$$

where $z_{ci} = \omega_{ci1}^{\mathrm{T}}[e_i, u_i, u_{-i}]$, ω_{ci1} and ω_{ci2} are the weight matrices, and $\Phi(x)$ is an activation function, that is, $\Phi(x) = \tanh(x) = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{\mathrm{e}^x + \mathrm{e}^{-x}}$.

Define the error function of this network as $e_{ci} = \hat{H}_i - H_i^*$. Because $H_i^* = 0$, we have $e_{ci} = (\nabla \hat{V}_i)^{\mathrm{T}} \dot{e}_i + r_i(e_i, u_i, u_{-i}, v_i, v_{-i})$.

Then, our goal is to minimize the loss function $E_{ci} = \frac{1}{2} e_{ci}^2$. Here, we adopt a gradient-based weight update rule. Thus,

$$\dot{\omega}_{ci} = -\beta_c \frac{\partial E_{ci}}{\partial \omega_{ci}} = -\beta_c \frac{\partial E_{ci}}{\partial e_{ci}} \frac{\partial e_{ci}}{\partial \hat{V}_i} \frac{\partial \hat{V}_i}{\partial \omega_{ci}}, \quad (14)$$

where β_c is a positive learning rate. Note that the weight update rule is applied to both ω_{ci1} and ω_{ci2}.

The actor-network of agent i uses its tracking error e_i as its input, and the output is $o_i = [u_i, v_i]$, which can be calculated by

$$o_i = \omega_{ai2}^{\mathrm{T}} \Phi(z_{ai}), \quad (15)$$

where $z_{ai} = \Phi(\omega_{ai1}^{\mathrm{T}} e_i)$.

Define the error function of the actor-network as $e_{ai} = \hat{V}_i - U_c$, where U_c is the designed cost-to-go objective. Generally, U_c is set to 0 to represent "success". Let the loss function be $E_{ai} = \frac{1}{2} e_{ai}^2$. Then, the weight update rule is adopted to minimize E_{ai}:

$$\dot{\omega}_{ai} = -\beta_a \frac{\partial E_{ai}}{\partial \omega_{ai}} = -\beta_a \frac{\partial E_{ai}}{\partial e_{ai}} \frac{\partial e_{ai}}{\partial o_i} \frac{\partial o_i}{\partial \omega_{ai}}, \quad (16)$$

where β_a is a positive learning rate.

The proposed results are illustrated via numerical simulation in Appendix B.

Conclusion. This study has addressed the containment control problem for MASs with unknown disturbances using a data-driven method. It has shown that the original containment control problem of MASs with unknown disturbances can be converted to the optimal containment control problem without disturbances.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61473061, 71503206, 61104104) and Program for New Century Excellent Talents in University (Grant No. NCET-13-0091).

Supporting information Appendixes A and B. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Wen G H, Zhao Y, Duan Z S, et al. Containment of higher-order multi-leader multi-agent systems: a dynamic output approach. IEEE Trans Autom Control, 2016, 61: 1135–1140
2 Lou Y C, Hong Y G. Target containment control of multi-agent systems with random switching interconnection topologies. Automatica, 2012, 48: 879–885
3 Meng Z Y, Ren W, You Z. Distributed finite-time attitude containment control for multiple rigid bodies. Automatica, 2010, 46: 2092–2099
4 Haghshenas H, Badamchizadeh M A, Baradarannia M. Containment control of heterogeneous linear multi-agent systems. Automatica, 2015, 54: 210–216
5 Peng Z N, Zhao Y Y, Hu J P, et al. Data-driven optimal tracking control of discrete-time multi-agent systems with two-stage policy iteration algorithm. Inf Sci, 2019, 481: 189–202
6 Peng Z N, Hu J P, Ghosh B K. Data-driven containment control of discrete-time multi-agent systems via value iteration. Sci China Inf Sci, 2020, 63: 189205