

博弈论框架下的二阶多智能体系统领导者选择算法

薛磊, 王庆领, 孙长银[†]

(东南大学 自动化学院, 江苏 南京 210096; 教育部 复杂工程系统测量与控制教育部重点实验室, 江苏 南京 210096)

摘要: 随着人工智能的发展, 多智能体系统中智能体的异质性、工作环境的复杂性、系统目标的多样性, 给多智能体系统动态性能的分析带来了挑战. 同时, 也催生了新的控制策略和优化方法. 博弈论作为一种研究社会系统中智能体决策过程的经典工具, 如今已被应用到了多智能体系统研究领域. 本文主要针对二阶多智能体系统编队过程中领导者选取的两类问题: 1) 选取 k 个领导者使系统误差达到最小; 2) 系统误差在一定范围内, 选取最小数量的领导者; 提出了一类超模博弈建模方法. 在建模过程中设计了各个智能体的效用函数与系统整体的目标函数, 使各个智能体在寻求各自效用函数最大化的过程中实现整体目标. 而后, 运用贪婪算法优化了智能体决策过程. 本文分析了所建模型的平衡点存在性和系统稳定性. 最后, 利用仿真实例对比说明了本文提出的基于超模博弈的二阶多智能体系统领导者选择算法的有效性.

关键词: 超模博弈; 多智能体系统; 领导者选取

中图分类号: TP273 文献标识码: A

Game theoretical approach for the leader selection of the second-order multi-agent system

XUE Lei, WANG Qing-ling, SUN Chang-yin[†]

(School of Automation, Southeast University, Nanjing Jiangsu 210096, China;
Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education,
Nanjing Jiangsu 210096, China)

Abstract: With the development of the artificial intelligence, the features of Multi-agent System (MAS), such as heterogeneous agents, complex environment, and objective variety, bring challenges for analysis the dynamics of the system. Therefore, new control policies and optimization methods are required to deal with the challenges. Game theory which is a classic tool for studying the decision making processes of the agents in social system is rising as a meaningful method for analyzing the dynamics of the MAS. Two kinds of leader selection problems which exist in the formation of second-order MAS are discussed. The supermodular game is introduced to model two kinds of leader selection problems, namely, the problem of selecting a fixed number of leaders in order to minimize the convergence error, as well as the problem of selecting the minimum-size set of leader agents to achieve a given bound on the convergence error. As to the designed supermodular game theoretical model, the utility functions and objective functions of the agents are designed. The global objective function can be achieved by maximizing utility function of the agents. Moreover, the greedy algorithm is proposed to optimize the decision making processes of the agents. The analytical properties of the designed model are also discussed. Simulation examples are provided to illustrate the effectiveness of the designed game theoretical model.

Key words: supermodular game; multi-agent system; leader selection

1 引言(Introduction)

多智能体系统的研究主要包括以下几个方面: 针对一致性控制^[1-3]、编队控制^[4-5]、以及分布式优化与

估计. 再者, 由于其具有广泛的应用基础, 比如传感器网络、机器人系统、工程系统等方面. 所以越来越多的学者致力于多智能体系统的研究. 例如, 在一致性

收稿日期: 2016-06-30; 录用日期: 2016-10-18.

[†]通信作者. E-mail: cysun@seu.edu.cn; Tel.: +86 13913036082.

本文责任编辑: 苏剑波.

国家自然科学基金项目(61503079, 61520106009, 61533008), 省自然科学基金项目(BK20150625), 江苏高校优势学科建设工程资助项目, 复杂工程系统测量与控制教育部重点实验室(MCCSE2015B02), 江苏省研究生创新基金项目(CXLX1309)资助.

Supported by National Natural Science Foundation of China (61503079, 61520106009, 61533008), National Natural Science Foundation of Province (BK20150625), a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions, the open fund of Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education (MCCSE2015B02) and Research Innovation Program for College Graduates of Jiangsu Province (CXLX1309).

领域,文献[6]研究了二阶多智能体系统达到一致性的充分必要条件.文献[7]讨论了二阶多智能体系统在固定拓扑和切换拓扑下的有限时间一致性.相对于一致性问题,多智能体编队问题在实际系统中具有更广泛的应用,比如无人机编队、机器人编队等方面.多智能体编队问题是指智能体通过彼此之间的信息交互形成预先设计的几何构型.在编队过程中,可以根据有无参照系分为两种:具有一般参照系的编队行为称为编队跟踪(formation tracking);没有一般参照系的编队行为称为编队形成(formation producing).对于编队跟踪来说,多智能体系统是由两种智能体组成,一种是领导者(leader),另一种是跟随者(follower).这种多智能体系统也叫作二阶领导者-跟随者系统(leader-follower system).在这种多智能体系统中,领导者的行为可以视作系统的控制输入,因为领导者的行为会对跟随者的状态(state)和速度(velocity)产生影响,从而引导整个系统的行为.近年来,很多学者在领导者-跟随者系统研究上,取得了很多卓越的研究成果.其中,文献[8]表明系统的性能会依赖于领导者的行为.文献[9]指出领导者的行为会作为控制输入影响跟随者的状态.因此,领导者的选择不仅会影响整个系统的收敛速度,还会影响跟随者之间的信息交互能力.研究领导者选择问题对提升系统性能和控制效果有着很大的帮助,对领导者选择问题的研究也逐渐开展起来.其中,文献[10]根据收敛误差最小化的方法来选择领导者,利用优化方法设计了一种领导者选择算法.文献[11]提出了一种矩阵优化策略来解决领导者选择问题.文献[12]介绍了一种多项式算法,通过运用这种多项式算法来寻找领导者集合,从而增加网络的可控性.已有的文献针对领导者选择问题的研究主要还是一阶多智能体系统,采用的方法多是经典的控制方法.因此,在这些研究成果的基础上,本文主要研究二阶多智能体系统的领导者选择问题,并提出了一种基于博弈论的针对二阶多智能体系统的领导者选择方法.

博弈论作为一种在竞争性的环境下研究个体之间相互作用与决策过程的工具,越来越多的被应用到工程领域,尤其是无线传感器网络领域.但是随着进化计算和人工智能的发展,很多学者开始运用传统控制方法以外的方式来解决实际问题^[13-14].在文献[15-17]中,博弈论作为优化控制方法的工具,应用到研究多智能体一致性问题的过程中.文献[14]结合了协调控制,强化学习,与博弈论方法,实现了多智能体在线编队.文献[17]提出了一种离散的多智能体动态博弈方法,来模拟智能体之间的相互作用.文献[18]研究了势博弈方法与协调控制的相关性,给博弈论在多智能体控制领域的应用带来了新的途径.博弈论作为研究协调控制问题的新方法,还处于起步阶段.因此,本

文在分析超模博弈模型特点的基础上,运用超模博弈解决二阶多智能体领导者选择问题.

本文所研究的二阶多智能体系统包含两类智能体:一类是普通点(跟随者),他们的决策过程取决于与相邻智能体之间的信息交互;另一类是特殊点(领导者),他们的决策过程取决于系统目标.这个设定可以应用到很多实际情况中,例如传感器网络,多无人机系统,以及多小车系统.例如多小车系统中,所有的小车都装有感应装置(这种装置可以与一定距离内的其他小车产生信息交互);作为领导者的小车还装有GPS定位装置.基于这样的系统,本文的主要工作包括:

1) 运用超模博弈对两类领导者选择问题进行建模.一类是在领导者数目固定的情况下,使得在编队过程中的收敛误差最小化;另一类问题是在收敛误差在一定范围内,如何选取最小数目的领导者.根据模型的特点,提出了贪婪算法来达到博弈论模型的纳什均衡点.

2) 本文以二阶分布式多智能体系统作为研究对象,通过与度启发式算法的对比仿真,来验证所提出方法的有效性.结果表明,所设计的超模博弈模型不仅在连通图中有较好的性能,而且在非连通图中也十分有效.

本文的组织结构如下:第1节为引言;第2节描述了两种存在于二阶多智能体编队过程中的领导者选择问题;第3节基于超模博弈,设计了两种领导者选择问题的超模博弈模型,同时也分析了模型的稳定性;第4节通过仿真验证了所设计模型的有效性;第5节是本文的结论.

注1 $\|f(\cdot)\|_2$ 表示 $f(\cdot)$ 的二范数, $|N|$ 表示集合 N 中所包含的元素数量

2 问题阐述(Problem statement)

本节主要包含两个部分,第1部分介绍了二阶多智能体系统;第2部分提出了两种领导者选择问题.

2.1 二阶多智能体系统(Second-order multi-agent system)

假设一个多智能体系统包含 $N = \{1, 2, \dots, n\}$ 个智能体,这 N 个智能体包括两个集合:领导者集合 N_L 与跟随者集合 N_F .由于多智能体系统的拓扑结构可以用图论来分析,本文利用图论的相关知识,假设邻接矩阵 $\Omega = [\omega_{ij}]$; $i, j = 1, \dots, n$ 用来表示智能体之间的连接关系.由于本文研究的多智能体系统的拓扑结构是无向图, ω_{ij} 就表示第 i 个智能体与第 j 个智能体可以进行通信.为了便于分析智能体与其相连的个体间的信息交互,引入 N_i 表示第 i 个智能体的邻居集合.本文分别讨论了在连通图与非连通图两种情况下的领导者选择问题.

对于二阶多智能体系统来说,它的动态特性可以

由以下方程组来表示:

$$\begin{cases} \dot{x}_i(t) = v_i(t), \\ \dot{v}_i(t) = u_i(t), \end{cases} \quad (1)$$

其中 x_i 与 v_i 分别表示第 i 个智能体的位置和速度.

2.2 领导者选择问题(Leader selection problem)

对于一阶系统来说, 智能体之间只交换他们的位置信息, 然而二阶多智能体则相对复杂. 当考虑控制问题的时候, 不仅要考虑智能体的位置信息, 还要顾及到智能体的速度信息. 因此, 要解决二阶多智能体系统在编队过程中的领导者选择问题, 必须满足以下 3 个目标:

1) 所有的智能体必须跟踪上外部输入的参考速度 v_r 和参考位置 x_r . 因此, 对于智能体 i 来说, $\|v_i - v_r\|$ 与 $\|x_i - x_r\|$ 需要收敛到 0.

2) 所有智能体必须保持设定的编队队形. 也就意味着智能体之间的相对位置要稳定. 假设利用矩阵 $D = \{d_{ij}\}$ 来表示智能体之间的相对位置, 其中 d_{ij} 表示智能体 i 和智能体 j 之间的相对距离.

3) 上述目标实现的过程中, 通过选取适当的领导者, 尽量减小过程中的收敛误差.

在上述 3 个目标中, 前两个是在解决编队问题的过程中, 需要解决的典型问题. 而第 3 个目标可以用来衡量领导者选择在优化编队过程中所产生的作用. 以上 3 个目标可以用方程表示为如下形式:

$$f(N_L) = \sum_{x_i \in N_L} [\|x_i - x_r\|_2 + \sum_{j \in N_i} \omega_{ij} \|v_i - v_j\|_2 + \sum_{j \in N_i} \omega_{ij} \|x_i - x_j - d_{ij}\|_2], \quad (2)$$

其中: $\omega = \omega_{ij}$ 是系统拓扑的邻接矩阵; N_L 是系统中的领导者集合; N_i 是可以与智能体 i 交互的个体集合; d_{ij} 表示智能体 i 和智能体 j 之间的相对距离; x_r 是系统外部输入的目标位置.

综上, 领导者选择问题可以转化为在编队过程中, 如何使收敛误差最小化的优化问题. 于是, 两类领导者选择问题可以定义如下:

定义 1 选取 k 个领导者, 使收敛误差最小化的过程, 可以定义为如下的形式:

$$\min f(N_L), \quad (3)$$

$$\text{s.t. } |N_L| \leq k, \quad (4)$$

其中: N_L 是系统中领导者集合, $|N_L|$ 为 N_L 中所包含的领导者数目.

定义 2 选取最少数量的领导者, 使收敛误差满足一定范围, 可以定义为如下的形式:

$$\min N_L, \quad (5)$$

$$\text{s.t. } f(N_L) \leq \alpha, \quad (6)$$

其中: N_L 是系统中领导者集合, $|N_L|$ 为 N_L 中所包含的领导者数目.

综上所述, 两类领导者选择问题可以转化为相对应的优化问题. 在优化过程中, 由于领导者行为可以积极的影响跟随者, 也就是说领导者在向系统目标移动的过程中, 可以通过信息交互将跟随者引向系统目标, 从而可以将领导者与跟随者之间的相互作用视作策略互补性的(strategic complementarities). 通过对超模博弈的了解, 发现策略互补性也正是超模博弈的特点. 然而, 就我们掌握的知识来说, 目前还没有尝试运用超模博弈来解决领导者选择问题的先例. 于是, 本文尝试运用超模博弈来解决领导者选择问题. 在下一个章节中, 将详细阐述针对两种领导者选择问题的超模博弈模型.

3 领导者选择问题的超模博弈模型(Supermodular game theoretical approach for leader selection problems)

在本节中, 首先介绍了超模博弈的相关概念, 随后讨论了两种领导者选择问题的超模博弈模型, 最后分析了模型的特性. 博弈论作为一种分析具有竞争性质现象的数学理论和方法, 对分析多智能体系统中个体间的相互作用有着独特的优势. 对于博弈论中的参与者来说, 他们的每一次决策都会拥有相应的回馈函数(payoff function). 这个回馈函数会受到其他参与者的影响, 从而影响下一轮博弈中的决策. 所有参与者决策的集合(joint action set)会依据系统的整体目标函数(global objective function)产生相应的数值. 因此, 构成博弈论模型的三要素如下: 智能体(agent)、每个智能体能够执行的行为(action)和这些行为产生的回馈函数(也称为utility function). 因此, 一个标准的博弈可以用 $G = (N, A, U)$ 来表示, 其中 N 是参与博弈过程的智能体; $A_i = \{a_1, a_2, \dots, a_k\} (i = 1, \dots, N)$ 是第 i 个智能体的行为集合; $U = \{u_1, u_2, \dots, u_N\}$, 其中 u_i 是第 i 个智能体的效用函数. 为了便于描述智能体的行为对系统整体目标函数产生的影响, 假设 (a_i, a_{-i}) 表示系统中所有智能体行为的集合, 其中 $a_{-i} = [a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_N]$ 表示除了第 i 个智能体以外, 其他所有智能体行为的集合. 以上是对于基本博弈模型的介绍, 对于超模博弈来说, 它的定义如下:

定义 3^[19] 对于一个博弈论模型 $G(N, A, U)$ 来说, 其中智能体分为两组 $N = \{N_m, N_n\}$, N_m 与 N_n 表示两种不同的智能体集合. 对于这两类智能体来说, 他们具有不同的行为集合 A_m, A_n , 与效用函数 U_m, U_n . 如果下列 3 项条件满足, 就称这个博弈模型为超模博弈:

- 1) A_m, A_n 是 R 的紧子集;
- 2) U_m, U_n 在 a_i 上是上半连续的, 在 a_{-i} 上是连续的;
- 3) U_m, U_n 在 A_m, A_n 上具有增大的差异.

引理 1^[17] 假设 $G(N, A, U)$ 是一个超模博弈模型, 其中智能体分为两组 $N = \{N_m, N_n\}$, N_m 与 N_n 表示两种不同的智能体集合. 对于这两类智能体来说, 他们具有不同的行为集合 A_m, A_n , 与效用函数 U_m, U_n . 使得

$$B_i(a_{-i}) = \arg \max_{a_{-i} \in A} P_i(a_i, a_{-i}), \quad (7)$$

那么

- 1) $B_i(a_{-i})$ 具有最大元素和最小元素, 分别表示为 $\bar{B}_i(a_{-i})$ 和 $\underline{B}_i(a_{-i})$.
- 2) 如果 $a'_{-i} \geq a_{-i}$, 那么 $\bar{B}_i(a'_{-i}) \geq \bar{B}_i(a_{-i})$, 并且 $\underline{B}_i(a'_{-i}) \geq \underline{B}_i(a_{-i})$.

超模博弈还具有如下的性质:

- 1) 对于超模博弈来说, 纯策略的纳什均衡是存在的;
- 2) 并且纯策略的纳什均衡是优势可解的(dominance solvable).

结合之前描述的领导者选择问题和超模博弈的特点, 它们本身存在着内部联系. 从系统组成的基本要素来看, 对于领导者选择问题来说, 智能体分为领导者集合与跟随者集合; 这两类智能体的行为集合和效用函数也可分别定义. 那么, 如果运用超模博弈来解决领导者选择问题, 有以下3个基本问题需要解决:

- 1) 领导者与跟随者的效用函数需要根据领导者选择问题的实际特点来定义, 同时, 效用函数还要满足超模博弈的特点;
- 2) 分析模型的稳定性, 明确模型的分析特性, 例如是否存在纳什均衡点;
- 3) 如果纳什均衡点存在, 需要设计适当的方法来优化智能体的决策过程, 从而达到纳什均衡点.

首先, 领导者选择问题的博弈论模型具有的基本要素为: 智能体 $N = \{N_L, N_F\}$, 其中: N_L 是领导者集合, N_F 是跟随者集合; 领导者与跟随者的行为集合分别表示为 A_F 与 A_L . 接下来, 针对智能体的不同特点与系统的目标函数, 本文设计了领导者与跟随者的效用函数如下:

$$\begin{cases} U_L = -\|x_i - x_0\|_2 - \sum_{j \in N_i} \omega_{ij} \|v_i - v_j\|_2 - \\ \sum_{j \in N_i} \omega_{ij} \|x_i - x_j - d\|_2, \\ U_F = -\sum_{j \in N_i} \|x_i - x_j - d_{ij}\|_2 - k \sum_{j \in N_i} \|v_i - v_j\|_2. \end{cases} \quad (8)$$

因此, 领导者选择问题的博弈论模型所包含的3个基本要素就设计完成了. 为了利用超模博弈的优势来解决领导者选择问题, 接下来需要证明所设计的博弈论模型为超模博弈.

定理 1 一个领导者选择问题的博弈论模型为 $G(N, A, U)$, 其中 $N = \{N_L, N_F\}$, N_L 为领导者集合, N_F 为跟随者集合; $A = \{A_L, A_F\}$, A_L 与 A_F 分别是领导者与跟随者的行为集合; $U = \{U_L, U_F\}$, U_L 与 U_F 分别为领导者与跟随者的效用函数. 如果下面3个性质满足, 则说明 G 是一个超模博弈:

- 1) A_L, A_F 是 R 的紧子集;
- 2) U_L, U_F 在 a_i 上是上半连续的, 在 a_{-i} 上是连续的;
- 3) U_L, U_F 在 A_L, A_F 上具有增长的差异.

此定理的证明可以通过以下步骤形成:

步骤 1 首先, 在实际系统中, 因为 A_L, A_F 为领导者与跟随者的行为集合, 所以 A_L 与 A_F 是有界闭集, 所以可以称 A_L, A_F 是 R 的紧子集;

步骤 2 其次, 对于任意的 $a^* \in A_F$ 与 $a_1, a_2 \in A_L$, 由公式(9)定义的 U_L, U_F , 并且根据博弈论智能体寻求个体利益最大化的过程中, U_L, U_F 在行为集合 A_L, A_F 上是递增的, 如果 $a_1 > a_2$, 有

$$\begin{cases} U_L(a_1 \vee a_2, a^*) \geq U_L(a_2, a^*), \\ U_L(a_1 \wedge a_2, a^*) \geq U_L(a_1, a^*), \\ U_L(a_1, a^*) + U_L(a_2, a^*) \leq \\ U_L(a_1 \vee a_2, a^*) + U_L(a_1 \wedge a_2, a^*). \end{cases} \quad (9)$$

如果 $a_1 < a_2$, 有

$$\begin{cases} U_L(a_1 \vee a_2, a^*) \geq U_L(a_1, a^*), \\ U_L(a_1 \wedge a_2, a^*) \geq U_L(a_2, a^*), \\ U_L(a_1, a^*) + U_L(a_2, a^*) \leq \\ U_L(a_1 \vee a_2, a^*) + U_L(a_1 \wedge a_2, a^*). \end{cases} \quad (10)$$

所以说 U_L, U_F 在 a_i 上是上半连续的, 在 a_{-i} 上是连续的;

步骤 3 对于任意的 $a_1, a_2 \in A_L$ 与 $b_1, b_2 \in A_F$, 如果有 $a_1 \geq a_2, b_1 \geq b_2$, 因为 U_L, U_F 如方程(9)所示, 存在

$$\begin{aligned} U_L(a_1, b_1) - U_L(a_1, b_2) &\geq U_L(a_2, b_1) - U_L(a_2, b_2), \\ U_F(a_1, b_1) - U_F(a_1, b_2) &\geq U_F(a_2, b_1) - U_F(a_2, b_2), \end{aligned}$$

所以, 可以说明 U_L, U_F 在 A_L, A_F 上具有增长的差异.

因此可以说明, 本文所设计的博弈模型为超模博弈模型.

根据定理1, 本文所设计的模型为超模博弈模型. 对于超模博弈来说, 纯策略纳什均衡点是一定存在的, 所以另一个需要解决的问题是寻找恰当的方法来优化智能体的决策过程来达到纳什均衡点. 因此, 本文

运用贪婪算法来优化智能体的决策过程. 在贪婪算法的每一步迭代过程中, 智能体依据总体贪婪所得来选择自己的下一步行动. 每一个行为就会得到相应的效用函数值 u_i^t . $L_i^T = \sum_{t=1}^T |u_i^t|$ 表示 a_i 在时刻 T 以内所得到的总体贪婪数值. 在选取下一步行为的时候, 通过比较 L_i^T 来选取具有最小值的行为作为下一次决策.

对于所设计的超模博弈模型来说, 系统整体目标函数可以通过智能体效用函数最大化的过程来实现. 而智能体追求自身利益最大化的过程, 也就是博弈模型追求纳什均衡点的过程. 因此当系统达到稳定, 也是博弈论模型达到纳什均衡的过程. 综上所述, 博弈论方法与传统控制方法的区别在于博弈论方法是通过优化智能体的决策过程, 让其根据自身所处的环境做出最优判断, 从而达到系统整体目标.

以上本文讨论了针对领导者选择问题所构建的超模博弈模型, 在下一小节中, 本文尝试运用超模博弈模型来解决两类具体的领导者选择问题.

3.1 选取 k 个领导者的超模博弈模型 (Supermodular game theoretical model for selecting k leaders)

在第 2 节中, 本文将编队过程中存在的领导者选择问题转化为一类优化问题, 其中, 选取 k 个领导者使收敛误差最小化的过程可以定义如下:

$$\min f(N_L), \tag{11}$$

$$\text{s.t. } |N_L| \leq k. \tag{12}$$

基于之前讨论的超模博弈模型, 针对选取 k 个领导者的超模博弈模型可以定义如下:

博弈论模型 $G(N, A, U)$, 其中 $N = \{N_L, N_F\}$, N_L 是领导者集合, N_F 是跟随者集合; $A = \{A_L, A_F\}$, A_L 与 A_F 分别是领导者与跟随者的行为集合; $U = \{U_L, U_F\}$, U_L 与 U_F 分别是领导者与跟随者的效用函数如方程(9)所示. 那么, 选取 k 个领导者的优化过程可以定义如下:

$$\begin{cases} \min f(N_L), \\ \text{s.t. } |N_L| \leq k, \\ f(N_L) = \\ \sum_{x_i \in N_L} [\|x_i - x_r\|_2 + \sum_{j \in N_i} \omega_{ij} \|v_i - v_j\|_2 + \\ \sum_{j \in N_i} \omega_{ij} \|x_i - x_j - d_{ij}\|_2]. \end{cases} \tag{13}$$

然而对于超模博弈模型来说, 不仅需要找到最优的领导者, 还要帮助智能体达到系统目标, 所以, 利用贪婪算法来逼近模型最优解的过程如下所示:

算法 1 选取 k 个领导者的算法.

输入: 领导者数目 k . 多智能体系统的拓扑结构 G ,

领导者的行为集合 $A_L = \{a_1, \dots, a_s\}$.

输出: 领导者集合 N_L .

初始化: $N_L = 0$, 当前领导者数目设置为 $m = 0$.

1) 如果 $m < k$, 那么根据以下方程选取领导者:

$$i \leftarrow \arg \max \{f(N_L) - f(N_L \cup \{i\})\},$$

$$N_L \leftarrow N_L \cup i,$$

$$m = m + 1.$$

如果 $m = k$, 表明领导者选取完毕, 进入下一步.

2) 如果 $i \in N_L$, 根据如下方程选取领导者最优行为

$$L_i^t = \sum_{T=1}^t U_i^T,$$

$$L_{\min}^{t-1} = \min L_j^{t-1}; a^{t-1} = \{i : L_j^{t-1} = L_{\min}^{t-1}\}.$$

最终智能体收敛于整体目标.

3.2 选取最少的领导者使系统误差满足一定范围的超模博弈模型 (Supermodular game theoretical model for selecting minimum leaders that achieves an error bound)

在本节中, 基于第 2 节中所介绍的选取最少领导者问题, 其定义如下:

$$\min N_L, \tag{14}$$

$$\text{s.t. } f(N_L) \leq \alpha. \tag{15}$$

结合之前讨论的超模博弈模型, 选取最少领导者的超模博弈模型可以定义如下:

博弈论模型 $G(N, A, U)$, 其中 $N = \{N_L, N_F\}$, N_L 是领导者集合, N_F 是跟随者集合; $A = \{A_L, A_F\}$, A_L 与 A_F 分别是领导者与跟随者的行为集合; $U = \{U_L, U_F\}$, U_L 与 U_F 分别是领导者与跟随者的效用函数如方程(9)所示. 那么, 选取 k 个领导者的优化过程可以定义如下:

$$\begin{cases} \min N_L, \\ \text{s.t. } f(N_L) \leq \alpha, \\ f(N_L) = \\ \sum_{x_i \in N_L} [\|x_i - x_r\|_2 + \sum_{j \in N_i} \omega_{ij} \|v_i - v_j\|_2 + \\ \sum_{j \in N_i} \omega_{ij} \|x_i - x_j - d_{ij}\|_2]. \end{cases} \tag{16}$$

基于上面描述的超模博弈模型, 利用贪婪算法寻找模型最优解的过程如下:

算法 2 选取最少领导者算法.

输入: 领导者数目 k . 多智能体系统的拓扑结构 G , 领导者的行为集合 $A_L = \{a_1, \dots, a_s\}$.

输出: 领导者集合 N_L .

初始化: $N_L = 0, \alpha$.

1) 如果 $f(N_L) > \alpha$, 那么利用以下方程选取领导

者:

$$\begin{aligned} i &\leftarrow \arg \max \{f(N_L) - f(N_L \cup \{i\})\}, \\ N_L &\leftarrow N_L \cup i. \end{aligned} \quad (17)$$

当k个领导者选取完毕,进入下一步.

2) 如果 $i \in N_L$, 则根据以下方程选取最优行为:

$$\begin{cases} L_i^t = \sum_{T=1}^t U_i^T, \\ L_{\min}^{t-1} = \min L_j^{t-1} \text{ a}^{t-1} = \{i : L_j^{t-1} = L_{\min}^{t-1}\}. \end{cases} \quad (18)$$

最终完成系统的整体目标.

由于选取最小数目的领导者的最优解与所要求的误差范围有关,那么最优解的范围与误差限的关系由以下定理给出:

定理 2 领导者集合 N_L 为选取最少领导者问题的最优解, N'_L 为贪婪算法寻找到的领导者集合, 那么

$$\frac{|N'_L|}{|N_L|} \leq 1 + \ln \frac{f_{\max}(N_L)}{\alpha}, \quad (19)$$

其中: α 为系统的最小收敛误差限; $f(\cdot)$ 为关于 N_L 的非下降次模函数.

此定理的证明,可以通过以下步骤形成:

步骤 1 首先,在文献[21]中,已经证明通过贪婪算法所得的 N'_L 与最优解 N_L 之间的关系为

$$\frac{|N'_L|}{|N_L|} \leq 1 + \ln \frac{f(N_L) - f(\phi)}{f(N_L) - f(N_L^{T-1})}, \quad (20)$$

其中: N_L 是最优的领导者点集, N_L^{T-1} 表示在倒数第2次迭代后所得到的领导者点集.

步骤 2 使 $f(N_S) = R_{\max} - R(N_S)$, 其中 $f(N_S)$ 是关于 N_L 非下降的次模函数.

$$\begin{aligned} \frac{|N'_L|}{|N_L|} &\leq \\ &1 + \ln \frac{f(N_L) - f(\phi)}{f(N_L) - f(N_L^{T-1})} = \\ &1 + \ln \frac{[R_{\max} - R(N_L)] - [R_{\max} - R(\phi)]}{[R_{\max} - R(N_L)] - [R_{\max} - R(N_L^{T-1})]} = \\ &1 + \ln \frac{-[R(N_L) - R(\phi)]}{-[R(N_L) - R(N_L^{T-1})]}, \end{aligned} \quad (21)$$

当 $R(N_L) = R_{\max}$ 时,也就是说当 N_L 为最优解时,

$$\begin{aligned} \frac{|N'_L|}{|N_L|} &= 1 + \ln \frac{-f(\phi)}{-f(N_L^{T-1})} = \\ &1 + \ln \frac{R_{\max}}{f(N_L^{T-1})}. \end{aligned} \quad (22)$$

由于 $f(N_L^{T-1}) \leq \alpha$,

$$\frac{|N'_L|}{|N_L|} \leq 1 + \ln \frac{f_{\max}(N_L)}{\alpha}, \quad (23)$$

因此,定理证明完毕.

综上所述,在本节中建立了两种领导者选择问题的超模博弈模型,通过贪婪算法来逼近问题的最优解.在下一节中,运用仿真实例来说明方法的有效性.

4 仿真实例(Simulation example)

本节中包含两个仿真实例来分别说明超模博弈模型在不同拓扑结构下针对两类领导者选择问题的有效性.考虑一个具有50个点的随机网络,这些点随机分布在 $1 \text{ cm} \times 1 \text{ cm}$ 的网格中.试验中假设每个智能体只与其相距 0.2 cm 范围内的其他智能体通信.为了验证超模博弈方法的有效性,仿真实验分别在连通图与非连通图中进行,同时引入了度启发式领导者选择算法 (degree heuristics algorithm) 与自稳定式领导者选择算法^[22](self-stabilizing algorithm) 作为对比.其中度启发式算法是通过比较拓扑结构中每个节点度的大小,选取度较大的点作为领导者.这样可以保证在选取领导者的过程中,每次选取的领导者都是对周围影响最大的节点.但是这种选取方法往往会忽视度相对较小的节点,例如在非连通图中,度较小的节点就可能接收不到外界的信息,从而无法达到系统目标.自稳定式领导者选择算法是通过寻找图的中心点和中心值来确定系统的领导者.

4.1 在连通图下的仿真实例 (Simulation example in connected graph)

随机网络在连通图下的拓扑结构如图1所示.

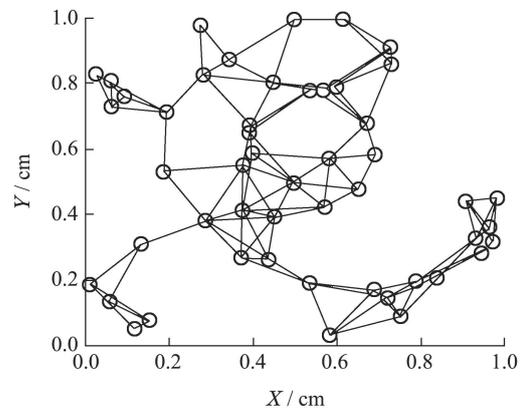


图 1 随机网络连通图

Fig. 1 Connected graph of random network

基于图1的拓扑结构,本文分别设定领导者数目为5与20,与两种其他领导者选择算法(度启发式算法与自稳定领导者选择算法)进行对比仿真.仿真结果如图2-6所示.

图2-5中的黑色点代表被选取作为领导者的点,通过对比可以看出,基于超模博弈模型选取的领导者较为均匀的分布在网络拓扑结构中,而利用度启发式算法选取的领导者相对集中在拓扑结构的某一区域中.通过图6和表1中的仿真结果表示在相同系统误差的情况下,相比于度启发式算法和自稳定领导者选择算法,超模博弈模型选取了更少的领导者.说明超模博

弈模型可以选取较少数量的领导者满足收敛误差限, 证明了超模博弈模型在处理选取最少领导者问题的有效性. 与此同时, 根据图6和表1, 在相同的领导者数目的情况下, 相比于其他两种领导者选择算法, 超模博弈模型具有更小的系统误差, 说明了超模博弈模型在选取 k 个领导者的情况下, 具有较小的误差.

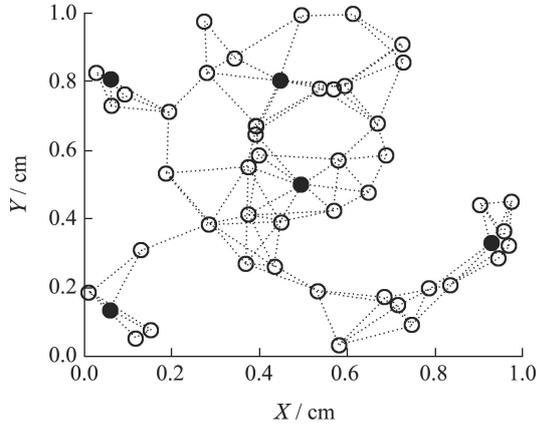


图 2 利用超模博弈方法选取 $k = 5$ 领导者

Fig. 2 Selecting $k = 5$ leaders via supermodular game theoretical method

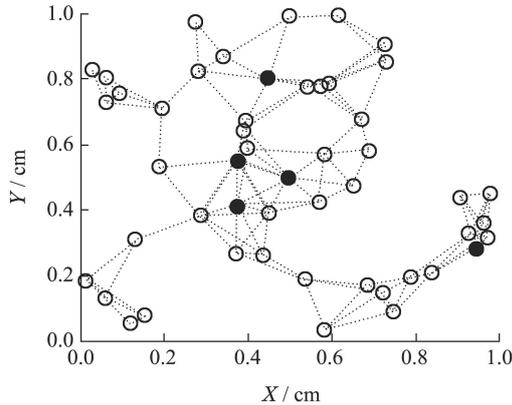


图 3 运用度启发式算法选取 $k = 5$ 领导者

Fig. 3 Selecting $k = 5$ leaders via degree heuristics algorithm

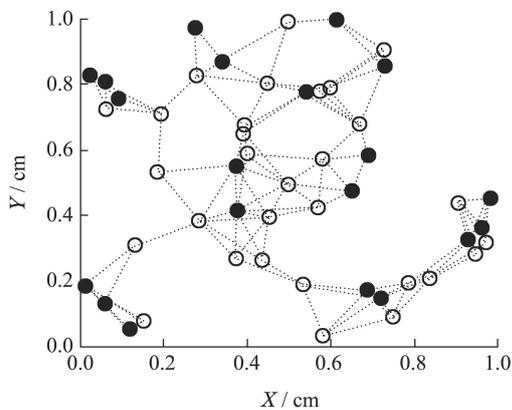


图 4 利用超模博弈方法选取 $k = 20$ 领导者

Fig. 4 Selecting $k = 20$ leaders via supermodular game theoretical method

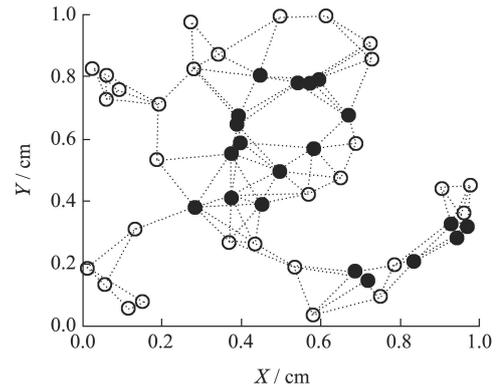


图 5 运用度启发式算法选取 $k = 20$ 领导者

Fig. 5 Selecting $k = 20$ leaders via degree heuristics algorithm

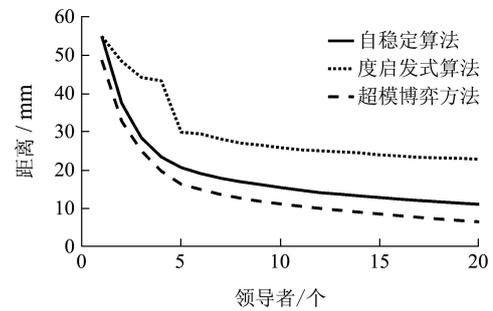


图 6 领导者数目的改变与收敛误差的关系

Fig. 6 The relationship between the number of leaders and convergence error

表 1 连通图下领导者数目与收敛误差的关系
Table 1 Relationships between number of leaders and convergence error under connected topology

N_L	自稳定领导者选择算法/mm	度启发式算法/mm	超模博弈模型算法/mm
1	54.98	54.98	48.73
2	37.51	48.45	32.82
3	28.43	44.22	25.03
4	23.44	43.33	19.73
5	20.68	29.91	16.33
6	19.07	29.45	14.91
7	17.88	28.11	13.61
8	16.98	27.07	12.67
9	16.23	26.55	11.84
10	15.46	25.89	11.18
11	14.77	25.32	10.57
12	14.09	25.04	10.01
13	13.69	24.78	9.49
14	13.25	24.51	8.99
15	12.84	23.97	8.53
16	12.44	23.69	8.09
17	12.07	23.33	7.66
18	11.73	23.20	7.24
19	11.39	23.09	6.85
20	11.08	22.86	6.51

以上仿真结果验证了超模博弈方法在连通图中的

有效性, 接下来将验证超模博弈方法在非连通图中的有效性. 随机网络在非连通图下的拓扑结构如图7所示.

4.2 在非连通图下的仿真实例(Simulation example in unconnected graph)

具有非连通图的网络拓扑结构如图7所示.

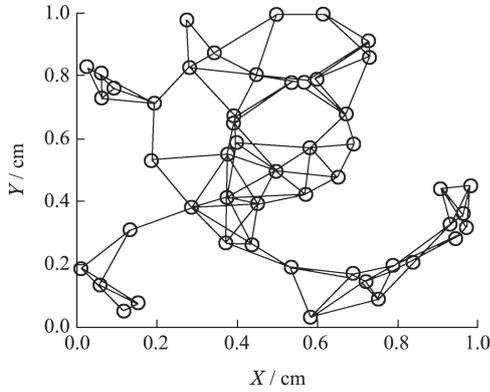


图7 随机网络非连通图

Fig. 7 Unconnected graph of random network

基于图7的拓扑结构, 本文分别设定领导者数目为5与20, 与其他两种领导者选择算法进行对比仿真. 仿真结果如图8-12与表2所示. 图8-11中黑色的点为被选取的领导者节点. 从图7可以看出, 在非连通图中存在着孤立的, 不连通的点集, 如果领导者选取不当, 会导致某些点集不能接收到控制信息, 从而影响控制结果. 如图9与图11所示, 度启发式领导者选取方法就错过了从图中两个孤立的点集中选取领导者, 而在其他的点集中选取了过多的领导者. 然而运用超模博弈方法在选取领导者的过程中, 则覆盖了这些孤立的点集. 从图12与表2的仿真结果可以看出, 在选取相同数目的领导者时, 超模博弈方法具有更小的误差, 说明了超模博弈方法在处理选取 k 个领导者问题时的有效性. 同时, 从图12和表2还可以发现, 当系统收敛误差相同的情况下, 超模博弈方法选取了较少的领导者节点, 说明了超模博弈方法在处理选取最少数目领导者问题时的有效性.

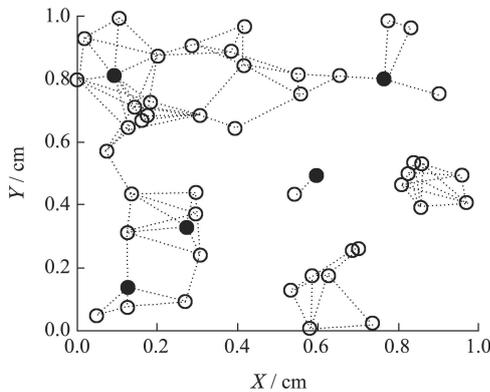


图8 利用超模博弈方法选取 $k = 5$ 领导者

Fig. 8 Selecting $k = 5$ leaders via supermodular game theoretical method

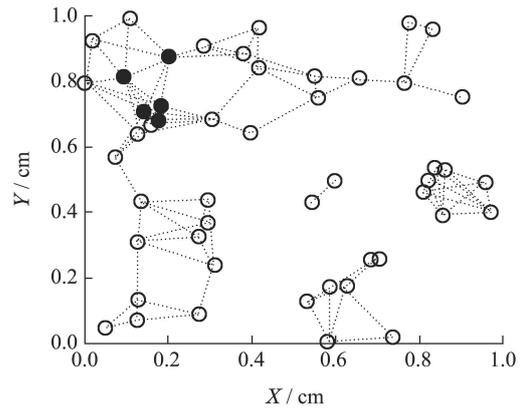


图9 运用度启发式算法选取 $k = 5$ 领导者

Fig. 9 Selecting $k = 5$ leaders via degree heuristics algorithm

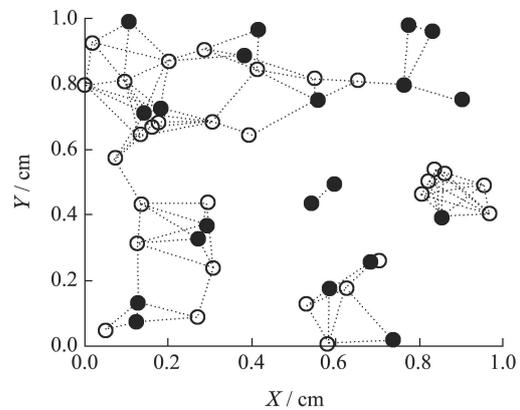


图10 利用超模博弈方法选取 $k = 20$ 领导者

Fig. 10 Selecting $k = 20$ leaders via supermodular game theoretical method

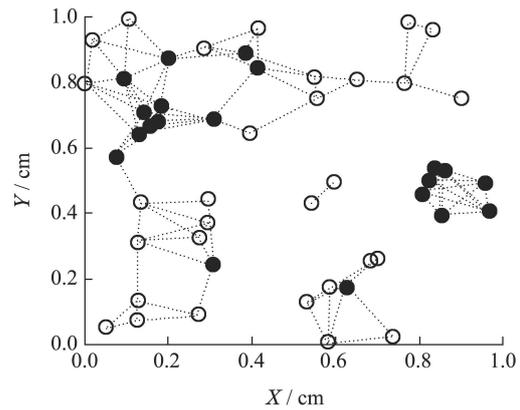


图11 运用度启发式算法选取 $k = 20$ 领导者

Fig. 11 Selecting $k = 20$ leaders via degree heuristics algorithm

在本节中, 通过与其他两种领导者选择算法(度启发式领导者选择算法与自稳定领导者选择算法)在连通图与非连通图中的表现相比较, 证明了超模博弈方法对解决两类领导者问题的有效性. 在选取 k 个领导者, 使误差最小的问题上, 相比于其他两种算法, 超模博弈模型具有更小的误差; 在选取最少数量的领导者,

使系统误差满足误差限的问题上, 相比于其他两种算法, 超模博弈选取了更少的领导者. 因此, 在文中所描述的随机网络中, 证明了超模博弈模型的有效性.

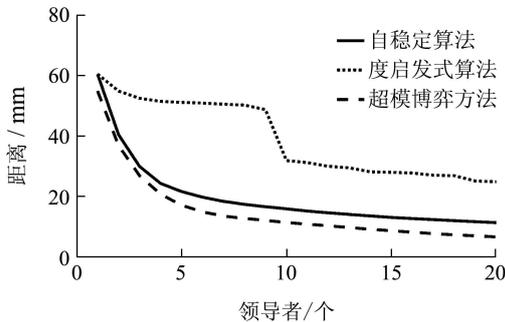


图 12 领导者数目的改变与收敛误差的关系

Fig. 12 The relationship between the number of leaders and convergence error

表 2 非连通图下领导者数目与收敛误差的关系

Table 2 Relationships between number of leaders and convergence error under unconnected topology

N_L	自稳定领导者 选择算法/mm	度启发式 算法/mm	超模博弈 模型算法/mm
1	60.39	60.44	54.84
2	40.44	54.86	36.81
3	29.94	52.45	26.89
4	24.33	51.46	20.78
5	21.66	51.13	17.15
6	19.78	50.89	14.91
7	18.41	50.50	13.54
8	17.41	50.21	12.78
9	16.65	48.75	12.09
10	15.92	31.89	11.44
11	15.21	31.21	10.86
12	14.63	30.01	10.31
13	14.07	29.46	9.75
14	13.59	28.16	9.21
15	13.12	28.01	8.70
16	12.73	27.76	8.21
17	12.36	27.04	7.74
18	12.02	26.88	7.37
19	11.69	25.13	7.01
20	11.38	24.87	6.65

5 结语(Conclusions)

针对两种领导者选择问题(选取 k 个领导者, 使系统的收敛误差最小; 选取最少数目的领导者, 使系统的收敛误差满足一定范围), 本文利用超模博弈对两类两道这选择问题进行建模, 采用了贪婪算法来优化智能体的决策过程, 通过与其他两种领导者选择算法相

比较, 说明了超模博弈在解决这两类问题上的有效性. 本文提出的方法具有以下特点:

1) 对于不同的领导者选择问题, 设计了超模博弈模型, 将两类领导者选择问题转化为了对应的优化问题, 便于利用超模博弈的方法进行研究;

2) 对于两种不同的领导者选择问题, 设计了超模博弈模型, 包括设计了系统的整体目标函数, 每个智能体的效用函数, 以及用于优化智能体决策过程的贪婪算法;

3) 通过与另外两种算法在随机网络中处理两类领导者选择的对比, 说明了超模博弈模型在处理两类领导者选择问题的有效性;

4) 本文结合了博弈论的相关内容解决了多智能体系统中存在的问题, 为后续将博弈论应用到分析更广泛的多智能体协调控制问题打下了基础.

参考文献(References):

- [1] YU W, CHEN G, REN W, et al. Distributed higher order consensus protocols in multi-agent dynamical systems [J]. *IEEE Transactions on Circuits & Systems I Regular Papers*, 2011, 58(8): 1924 – 1932.
- [2] XU X, LIU L, FENG G. Consensus of single integrator multi-agent systems with directed topology and communication delays [J]. *Control Theory and Technology*, 2016, 14(1): 21 – 27.
- [3] WANG L, XIE G. Consensus control for a class of networks of dynamic agents [J]. *International Journal of Robust and Nonlinear Control*, 2007, 17(10/11): 941 – 959.
- [4] ZHAO L, MA D. Circle formation control for multi-agent systems with a leader [J]. *Control Theory and Technology*, 2015, 12(1): 82 – 88.
- [5] PEI H, CHEN S, LAI Q. A local flocking algorithm of multi-agent dynamic systems [J]. *International Journal of Control*, 2015, 88(11): 1 – 23.
- [6] YU W, CHEN G, CAO M. Some necessary and sufficient conditions for second-order consensus in multi-agent dynamical systems [J]. *Automatica*, 2010, 46(3): 1089 – 1095.
- [7] LIN P, JIA Y. Consensus of a class of second-order multi-agent systems with time-delay and jointly-connected topologies consensus problems in networks of agents with switching topology and time-delays [J]. *IEEE Transactions on Automatic Control*, 2010, 55(3): 778 – 784.
- [8] LIN F, FARDAD M, JOVANOVIĆ M. Optimal control of vehicular formations with nearest neighbor interactions [J]. *IEEE Transactions on Automatic Control*, 2012, 57(9): 2203 – 2218.
- [9] SWAROOP D, HEDRICK J. String stability of interconnected systems [J]. *IEEE Transactions on Automatic Control*, 1996, 41(3): 349 – 357.
- [10] CLARK A, BUSHNELL L, POOVENDRAN R. A supermodular optimization framework for leader selection under link noise in linear multi-agent systems [J]. *IEEE Transactions on Automatic Control*, 2012, 59(2): 283 – 296.
- [11] LIU B, CHU T, WANG L, XIE G. Controllability of a leader-follower dynamic network with switching topology [J]. *IEEE Transactions on Automatic Control*, 2008, 53(4): 1009 – 1013.

- [12] LIU Y, SLOTINE J, BARABASI A. Controllability of complex networks [J]. *Nature*, 2011, 473(7346): 167 – 173.
- [13] XUE L, SUN C, MU C, et al. A RBF neural network learning algorithm based on NCP SO [C] // *The 32nd Chinese Control Conference (CCC)*. Hefei: IEEE, 2013: 3294 – 3299.
- [14] MU C, SUN C, XUE L. Design of fast sliding mode control on air-breathing hypersonic vehicles [C] // *The 32nd Chinese Control Conference (CCC)*. Hefei: IEEE, 2013: 805 – 810.
- [15] SEMSAR-KAZEROONI E, KHORASANI K. Multi-agent team cooperation: a game theory approach [J]. *Automatica*, 2009, 45(1): 2205 – 2213.
- [16] VAMVOUDAKIS K, LEWIS F, HUDAS G. Multi-agent differential graphical games: online adaptive learning solution for synchronization with optimality [J]. *Automatica*, 2012, 48(1): 1598 – 1611.
- [17] ABOUHEAF M, LEWIS F, VAMVOUDAKIS K, et al. Multi-agent discrete-time graphical games and reinforcement learning solutions [J]. *Automatica*, 2014, 50(1): 3038 – 3053.
- [18] MARDEN J, ARSLAN G, SHAMMA J. Cooperative control and potential games [J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2009, 39(6): 1393 – 1407.
- [19] BRANZEI R, MALLOZZI L, TIJS S. Supermodular games and potential games [J]. *Journal of Mathematical Economics*, 2003, 39(1/2): 1009 – 1013.
- [20] LIU B, CHU T, WANG L, et al. Controllability of a leader-follower dynamic network with switching topology [J]. *IEEE Transactions on Automatic Control*, 2008, 53(4): 1009 – 1013.
- [21] WOLSEY L, NEMHAUSER G. *Integer and Combinatorial Optimization* [M]. New York: Wiley-Interscience, 1999: 110 – 115.
- [22] Patterson S. In-network leader selection for acyclic graphs [C] // *American Control Conference (ACC)*. Hefei: IEEE, 2015: 329 – 334.

作者简介:

薛磊 (1986–), 男, 博士研究生, 目前研究方向为博弈论与多智能体系统, E-mail: rayxue@seu.edu.cn;

王庆领 (1982–), 男, 讲师, 目前从事多智能体系统及受限控制方面的研究、承担包括国家自然科学基金青年基金、江苏省自然科学基金等多个项目的研究工作, E-mail: csuwql@gmail.com;

孙长银 (1975–), 男, 教授, 博士生导师, 目前从事模式识别与智能系统控制领域研究, 正在承担包括国家重大研究计划重点支持项目、重点国际合作项目等国家级课题研究工作, E-mail: cysun@seu.edu.cn.