



# 多颗微小卫星接管失效航天器姿态运动的微分博弈学习控制

韩楠<sup>1,2</sup>, 罗建军<sup>1,2\*</sup>, 柴源<sup>1,2</sup>

1. 西北工业大学深圳研究院, 深圳 518057

2. 西北工业大学航天学院航天飞行动力学技术重点实验室, 西安 710072

\* 通信作者. E-mail: jjluo@nwpu.edu.cn

收稿日期: 2019-02-28; 修回日期: 2019-05-24; 接受日期: 2019-06-05; 网络出版日期: 2020-04-08

深圳科创委基金项目 (批准号: JCYJ20180508151938535)、国家自然科学基金重大项目 (批准号: 61690210, 61690211) 和西北工业大学博士论文创新基金项目 (批准号: CX201803) 资助

**摘要** 对空间中由于燃料耗尽或执行机构故障而失效的航天器进行姿态接管控制能够实现其上高价值有效载荷的再利用, 多个微小卫星通过协同, 能够以较低的成本对其实施姿态接管控制. 微分博弈研究了多个体的最优决策问题, 其中各个体通过局部性能指标函数的优化获得控制策略, 并实现博弈事先设定的全局目标. 本文将失效航天器的姿态接管控制问题转换为多微小卫星的微分博弈问题进行研究. 首先建立了多微小卫星的微分博弈模型, 设计了微小卫星的性能指标函数, 实现了对微小卫星微分博弈问题的数学描述. 之后给出了微小卫星微分博弈的 HJ (Hamilton-Jacobi) 方程, 并通过基于单神经网络的策略迭代算法进行了耦合 HJ 方程解的学习, 实现了对微小卫星博弈均衡策略的逼近. 最后进行了微小卫星微分博弈学习控制方法有效性的数值仿真验证, 仿真结果表明, 微小卫星通过学习到的近似博弈均衡控制策略, 实现了失效航天器姿态接管控制的全局目标.

**关键词** 姿态控制, 接管控制, 微分博弈, 博弈控制, 策略迭代

## 1 引言

空间中的失效航天器由于燃料耗尽或执行机构故障而处于不受控的自由漂浮状态, 无法进行有效的姿态定向以继续工作, 造成了其上高价值有效载荷的浪费. 对失效航天器实施姿态接管控制能够延长其在轨寿命, 这为合理处置这些失效航天器提供了一种新思路<sup>[1]</sup>. 与重新发射一颗新的替代卫星相比, 以姿态接管控制的方式实现对失效航天器上高价值有效载荷的再利用, 不仅能够降低空间任务的研发成本, 还能够有效缩短其研发周期.

**引用格式:** 韩楠, 罗建军, 柴源. 多颗微小卫星接管失效航天器姿态运动的微分博弈学习控制. 中国科学: 信息科学, 2020, 50: 588–602, doi: 10.1360/N112019-00049

Han N, Luo J J, Chai Y. Differential game learning approach for multiple microsattellites takeover of the attitude movement of failed spacecraft (in Chinese). Sci Sin Inform, 2020, 50: 588–602, doi: 10.1360/N112019-00049

接管控制是指服务航天器与目标航天器结合后,利用自身的执行机构接管目标的姿态和轨道控制功能,以实现目标航天器姿态和轨道的精确控制<sup>[2]</sup>.文献[2,3]针对基于空间机械臂的失效航天器姿态接管控制问题进行了研究.空间机械臂为实施失效航天器的姿态接管控制提供了一种有效的途径,但难以对失效航天器进行长期姿态调整,以满足高价值有效载荷再利用任务所提出的姿态控制要求.另外,针对不同任务定制相应的空间机器人系统将带来时间和成本的巨大消耗<sup>[4]</sup>.近年来,微小卫星技术得到了快速发展,以 DARPA (Defense Advanced Research Projects Agency, 美国国防部高级研究计划局) 提出的凤凰计划项目<sup>[5]</sup>和 DLR (German Aerospace Center, 德国宇航中心) 提出的 iBOSS 项目<sup>[6]</sup>为代表的任务,旨在通过发射不同类型的微小卫星,实现空间资源的再利用和新卫星的在轨组装.凤凰计划项目设想通过发射一些微小卫星与空间中失效航天器上的天线进行重新装配,来实现空间昂贵部件的回收利用,以降低空间设备的研发成本和发射成本.这一设想为通过微小卫星实现失效航天器的姿态接管控制提供了借鉴.微小卫星发射成本低、研制周期短、机动灵活,能够通过互相协同的方式,以较低的成本、较快的速度实现对失效航天器的长期姿态调整,进而实现对失效航天器上高价值有效载荷的再利用.

基于多微小卫星协同的失效航天器姿态接管控制方法涉及到多微小卫星的协同控制问题.过去 20 年,分布式协同控制问题在编队和集群飞行任务中得到了广泛关注.传统的分布式协同控制方法能够通过局部信息的感知实现多微小卫星的长期在轨飞行,但难以通过多微小卫星协同实现失效航天器的姿态接管控制.针对此类问题,文献[4]通过对期望控制量进行分布式分配实现了失效航天器的接管控制,但需要不断进行的控制分配增加了计算负担.近年来,随着进化计算和人工智能的发展,很多学者开始运用智能算法来解决实际问题<sup>[7]</sup>.微分博弈研究了多个体的最优决策问题,其中的每一个体通过局部性能指标函数的优化计算各自的控制策略,并能够在此基础上实现博弈事先设定的全局目标,近年来在多智能体控制领域得到了越来越多的关注.文献[8]通过微分博弈方法研究了无人机编队的控制问题,文献[9,10]将微分博弈方法应用到多智能体一致性问题的研究中,文献[11]研究了基于认知博弈制导的无人机自主防碰撞控制问题,文献[12,13]则研究了追逃博弈问题.

本文将基于多微小卫星协同的失效航天器姿态接管控制问题转换为多微小卫星的微分博弈问题进行处理,通过  $N$  人微分博弈问题中的博弈均衡方法<sup>[14~16]</sup>来求解微小卫星的 Nash 均衡策略,以实现微小卫星的自主决策,能够在避免进行控制分配的情况下实现对失效航天器的姿态接管控制,主要工作包括:

(1) 针对多微小卫星博弈接管失效航天器的姿态控制问题,建立了多微小卫星的微分博弈模型,设计了微小卫星的性能指标函数,微小卫星通过独立优化性能指标函数的方式计算控制策略,以实现失效航天器的姿态接管控制.

(2) 通过基于单神经网络的策略迭代算法逼近微小卫星的 Nash 均衡策略,并通过 Lyapunov 方法证明了权值矢量估计误差的一致最终有界性 (uniform ultimate boundedness, UUB).

本文的组织结构如下:第 2 节对本文所研究的问题进行了描述;第 3 节建立了多微小卫星微分博弈的数学模型;第 4 节给出了基于策略迭代算法的博弈均衡策略求解方法;第 5 节通过数值仿真验证了失效航天器姿态接管的微分博弈控制方法的有效性;第 6 节给出了本文的结论.

文中所涉及到的数学符号为: $A > 0$  表示  $A$  为正定矩阵, $A \geq 0$  表示  $A$  为半正定矩阵. $I_n$  表示  $n \times n$  的单位矩阵. $0_{m \times n}$  表示  $m \times n$  的全零矩阵. $1_n$  表示  $n$  维全 1 列向量. $a^\times$  表示矢量  $a \in \mathbb{R}^3$  的反对称矩阵,定义为  $a^\times = [0, -a_3, a_2; a_3, 0, -a_1; -a_2, a_1, 0] \in \mathbb{R}^{3 \times 3}$ .  $A^T, a^T$  分别表示矩阵  $A$ , 矢量  $a$  的转置. $\mathbb{N} := \{1, 2, \dots, N\}$  表示参与博弈的  $N$  颗微小卫星的集合.

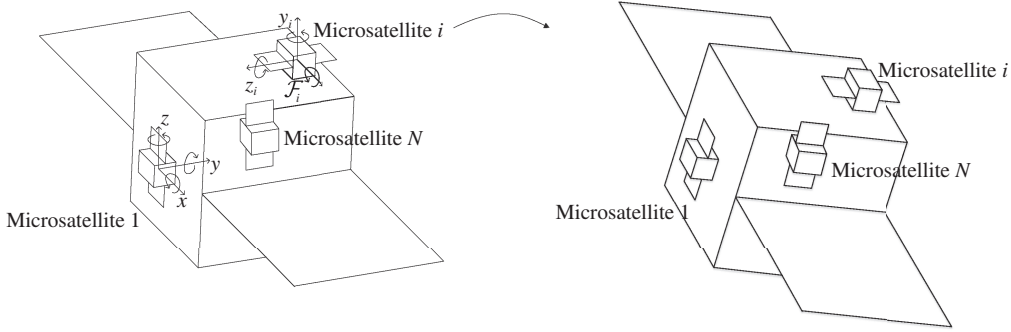


图 1 姿态接管控制示意图

Figure 1 Sketch of attitude takeover control

## 2 问题描述

为了实现对失效航天器的姿态接管控制, 微小卫星需要通过自身的执行机构为失效航天器提供控制力矩. 姿态接管控制示意图如图 1 所示. 假设:

- (1) 失效航天器和微小卫星均可被视为刚体.
- (2) 微小卫星之间的相对方位保持不变.
- (3) 失效航天器姿态接管控制所需的控制力矩只由微小卫星提供.

本文通过微分博弈方法来处理失效航天器的姿态接管控制问题. 微小卫星通过独立优化各自的性能指标函数来计算控制策略, 并在此基础上完成对失效航天器的姿态接管控制.

本文所涉及的坐标系定义如下:

(1) 微小卫星  $i$  ( $i \in \mathbb{N}$ ) 本体坐标系  $O_i x_i y_i z_i$ ,  $O_i$  表示微小卫星  $i$  的质心,  $x_i, y_i, z_i$  表示微小卫星  $i$  的 3 个体轴, 微小卫星  $i$  沿此三体轴方向产生控制力矩.

(2) 参考坐标系  $Oxyz$ , 参考坐标系用于描述各微小卫星本体坐标系的方位, 本文选取微小卫星 1 的本体坐标系作为参考坐标系, 其余微小卫星相对于微小卫星 1 的方位可由其各自本体坐标系相对于参考坐标系的转换矩阵来描述.

## 3 多微小卫星博弈数学建模

### 3.1 姿态运动学模型

由于微小卫星和失效航天器均假设是刚体, 因此本文通过刚体姿态运动方程来描述姿态接管控制任务中组合体的姿态运动<sup>[17]</sup>. 修正罗德里格斯参数 (modified Rodrigues parameters, MRPs) 提供了一种有效描述刚体转动的方式, 其通过刚体转动的转轴和转角来给出. 假设  $e$  为单位矢量, 有一刚体绕  $e$  轴转过  $\phi$  角, 则描述该转动的 MRPs 为  $\sigma = e \tan(\frac{\phi}{4}) \in \mathbb{R}^3$ , 相应的姿态运动学方程可通过  $\sigma$  的变化率给出:

$$\dot{\sigma} = G(\sigma)\omega, \quad (1)$$

其中  $\omega \in \mathbb{R}^3$  为刚体的姿态角速度,  $G(\sigma)$  由下式给出:

$$G(\sigma) = \frac{1}{4}[(1 - \sigma^T \sigma)I_3 + 2\sigma^\times + 2\sigma\sigma^T]. \quad (2)$$

式 (1) 即给出了组合体的姿态运动学模型.

### 3.2 姿态动力学模型

由于组合体姿态机动所需的控制力矩完全由微小卫星提供, 因此, 组合体的姿态动力学方程为

$$\dot{\omega} = -J^{-1}\omega \times (J\omega) + J^{-1}\sum_{j=1}^N C_j u_j, \quad (3)$$

其中  $J \in \mathbb{R}^{3 \times 3}$  为组合体的转动惯量矩阵,  $C_j \in \mathbb{R}^{3 \times 3}$  为微小卫星  $j$  本体坐标系到参考坐标系的转换矩阵,  $u_j \in \mathbb{R}^3$  为微小卫星  $j$  的控制策略 (控制力矩). 式 (3) 中, 微小卫星  $j$  通过调整  $u_j$  可直接控制组合体角速度  $\omega$  的变化, 并可通过  $\omega$  的变化来间接地控制式 (1) 中  $\sigma$  的变化. 因此, 微小卫星通过调整  $u_j$  ( $j \in \mathbb{N}$ ) 即可实现对组合体姿态状态量  $\sigma$  以及  $\omega$  的调整, 进而实现对组合体姿态运动的控制.

### 3.3 微小卫星微分博弈模型

根据式 (1) 和 (3), 可得组合体姿态运动方程为

$$\dot{x} = f(x) + \sum_{j=1}^N g_j u_j, \quad (4)$$

其中  $x = [\sigma^T, \omega^T]^T \in \mathbb{R}^6$ , 且

$$f(x) = \begin{bmatrix} G(\sigma)\omega \\ -J^{-1}\omega \times (J\omega) \end{bmatrix} \in \mathbb{R}^6, \quad (5a)$$

$$g_j = \begin{bmatrix} 0_{3 \times 3} \\ J^{-1}C_j \end{bmatrix} \in \mathbb{R}^{6 \times 3}, \quad (5b)$$

各颗微小卫星通过优化如下的性能指标函数来获得控制策略:

$$\begin{aligned} J_i(x_0, u_1, \dots, u_N) &= \int_0^\infty \left( x^T Q_i x + \sum_{j=1}^N u_j^T R_{ij} u_j \right) dt \\ &\triangleq \int_0^\infty r_i(x, u_1, \dots, u_N) dt, \quad i \in \mathbb{N}, \end{aligned} \quad (6)$$

其中  $Q_i > 0$ ,  $R_{ii} > 0$ ,  $R_{ij} \geq 0$  均为对称矩阵.

多微小卫星的博弈可通过式 (4) 与 (6) 来定义, 由于各微小卫星性能指标函数之和不为零 (非零和博弈), 且系统运动方程 (4) 为一微分方程, 因此, 该博弈实际上是一个非零和微分博弈. 此外, 微小卫星性能指标函数式 (6) 的被积函数中, 第 1 项为组合体姿态状态量  $x$  的二次函数, 优化该项的积分意味着进行组合体姿态的镇定, 第 2 项为各微小卫星控制量二次函数的加权组合, 优化该项的积分意味着优化各微小卫星的控制能量消耗, 因此, 当各微小卫星进行式 (6) 的优化时, 便可在尽可能最小化微小卫星控制能量消耗的情况下使组合体的姿态状态量  $x$  趋于稳定, 进而实现对失效航天器的姿态接管控制. 本文通过求解多微小卫星非零和微分博弈的 Nash 均衡策略来计算各微小卫星的控制策略.

**定义1** (可行控制策略集) 反馈控制策略集  $u(x) = \{u_1(x), \dots, u_N(x)\}$  在紧集  $\Omega$  上相对于式 (6) 定义为可行的, 记作  $u_i(x) \in \Psi(\Omega)$ , 如果  $u_i(x)$  在  $\Omega$  上连续,  $u_i(0) = 0$ ,  $u(x)$  在  $\Omega$  上稳定式 (4), 且式 (6) 对于任意的  $x_0 \in \Omega$  是有限的.

对应于可行控制策略集  $u_i(x) \in \Psi(\Omega)$  的值函数为

$$\begin{aligned} V_i(x(t), u_1, \dots, u_N) &= \int_t^\infty \left( x^T Q_i x + \sum_{j=1}^N u_j^T R_{ij} u_j \right) d\tau \\ &= \int_t^\infty r_i(x, u_1, \dots, u_N) d\tau, \quad i \in \mathbb{N}. \end{aligned} \quad (7)$$

**定义2** (Nash 均衡) 对于由式 (4) 与 (6) 所定义的微分博弈来说, 反馈控制策略集  $\{u_1^*, \dots, u_N^*\}$  构成微分博弈的 Nash 均衡, 如果下面的  $N$  个不等式均成立:

$$J_i(u_1^*, \dots, u_i^*, \dots, u_N^*) \leq J_i(u_1^*, \dots, u_i, \dots, u_N^*), \quad i \in \mathbb{N}. \quad (8)$$

$N$  颗微小卫星 Nash 均衡求解问题可描述为

$$\begin{aligned} V_i^*(x(t)) &= \min_{u_i \in \Psi(\Omega)} \int_t^\infty \left( x^T Q_i x + \sum_{j=1}^N u_j^T R_{ij} u_j \right) d\tau \\ &= \min_{u_i \in \Psi(\Omega)} \int_t^\infty r_i(x, u_1, \dots, u_N) d\tau, \quad i \in \mathbb{N}. \end{aligned} \quad (9)$$

## 4 多微小卫星博弈均衡

### 4.1 博弈均衡策略显式表达式

根据最优性原理, 式 (7) 可等价为

$$\begin{aligned} V_i(x(t), u_1, \dots, u_N) &= \int_t^\infty r_i(x, u_1, \dots, u_N) d\tau \\ &= \int_{t+\Delta t}^\infty r_i(x, u_1, \dots, u_N) d\tau + \int_t^{t+\Delta t} r_i(x, u_1, \dots, u_N) d\tau \\ &= V_i(x(t+\Delta t), u_1, \dots, u_N) + \int_t^{t+\Delta t} r_i(x, u_1, \dots, u_N) d\tau, \quad i \in \mathbb{N}, \end{aligned} \quad (10)$$

将  $V_i(x(t+\Delta t), u_1, \dots, u_N)$  通过泰勒 (Taylor) 级数展开, 可得

$$V_i(x(t+\Delta t), u_1, \dots, u_N) = V_i(x(t), u_1, \dots, u_N) + \frac{\partial V_i}{\partial x} \dot{x} \Delta t + \frac{\partial V_i}{\partial t} \Delta t + o(\Delta t^2), \quad i \in \mathbb{N}, \quad (11)$$

其中  $o(\Delta t^2)$  为二次及二次以上各项.

由于式 (4) 中姿态运动方程定常, 因此  $\partial V_i / \partial t = 0$ . 将式 (11) 代入式 (10) 中, 取  $\Delta t \rightarrow 0$ , 并以  $\Delta t$  除之, 可得

$$0 = r_i(x, u_1, \dots, u_N) + (\nabla V_i)^T \left( f(x) + \sum_{j=1}^N g_j u_j \right), \quad i \in \mathbb{N}, \quad (12)$$

其中  $\nabla V_i = \partial V_i / \partial x$ ,  $V_i(0) = 0$ . 式 (12) 给出了值函数 (7) 的微分等价. 给定一组控制量  $\{u_1, \dots, u_N\}$  时, 便可通过求解式 (12) 计算出对应于该组控制量的值函数<sup>[14]</sup>.

定义哈密尔顿 (Hamilton) 函数如下:

$$H_i(x, \nabla V_i, u_1, \dots, u_N) = r_i(x, u_1, \dots, u_N) + (\nabla V_i)^T \left( f(x) + \sum_{j=1}^N g_j u_j \right), \quad i \in \mathbb{N}, \quad (13)$$

对应于最优值函数 (9) 的反馈控制策略为

$$\frac{\partial H_i}{\partial u_i} = 0 \Rightarrow u_i^* = -\frac{1}{2} R_{ii}^{-1} g_i^T \nabla V_i^*, \quad i \in \mathbb{N}, \quad (14)$$

式 (14) 即为微小卫星 Nash 均衡策略的显式表达式. 将式 (14) 代入式 (12) 中, 可得  $N$  个耦合的 HJ 方程为

$$\begin{aligned} 0 = & x^T Q_i x + \frac{1}{4} \sum_{j=1}^N (\nabla V_j^*)^T g_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} g_j^T \nabla V_j^* \\ & + (\nabla V_i^*)^T f(x) - \frac{1}{2} (\nabla V_i^*)^T \sum_{j=1}^N g_j R_{jj}^{-1} g_j^T \nabla V_j^*, \quad i \in \mathbb{N}. \end{aligned} \quad (15)$$

求解式 (15) 中  $N$  个耦合的 HJ 方程, 可得  $\nabla V_i^*$ , 之后即可将其代入式 (14) 中求得微小卫星的 Nash 均衡策略. 然而, HJ 方程的非线性特性使得  $\nabla V_i^*$  的解析解很难求得, 为此, 可通过学习算法逼近其数值解.

## 4.2 策略迭代算法

策略迭代算法<sup>[14]</sup> 是一种强化学习算法, 包含策略评估和策略改进两个环节, 具体执行步骤如下.

第 1 步: 设定计算精度  $\delta > 0$ , 令迭代次数  $k = 0$ , 且  $V_i^0(x) = 0$ , 选择初始可行控制策略集  $u^0 = \{u_1^0, \dots, u_N^0\}$ .

第 2 步: 给定控制策略集  $u^k = \{u_1^k, \dots, u_N^k\}$ , 通过求解下述  $N$  个方程计算  $\{V_1^{k+1}, \dots, V_N^{k+1}\}$ :

$$0 = r_i(x, u_1^k, \dots, u_N^k) + (\nabla V_i^{k+1})^T \left( f(x) + \sum_{j=1}^N g_j u_j^k \right), \quad i \in \mathbb{N}. \quad (16)$$

第 3 步: 通过下式进行控制策略的更新:

$$\begin{aligned} u_i^{k+1} &= \arg \min_{u_i \in \Psi(\Omega)} H_i(x, \nabla V_i^{k+1}, u_1, \dots, u_N) \\ &= -\frac{1}{2} R_{ii}^{-1} g_i^T \nabla V_i^{k+1}, \quad i \in \mathbb{N}. \end{aligned} \quad (17)$$

第 4 步: 如果  $\max[\|V_1^{k+1} - V_1^k\| \dots \|V_N^{k+1} - V_N^k\|] \leq \delta$ , 迭代停止, 得到近似最优控制策略集  $u^{k+1} = \{u_1^{k+1}, \dots, u_N^{k+1}\}$ , 否则, 令  $k = k + 1$ , 返回第 2 步.

策略迭代算法的收敛性证明可参见文献 [15].

### 4.3 博弈均衡策略逼近

在紧集  $\Omega$  上, 微小卫星  $i$  的最优值函数可以用如下评价神经网络近似:

$$V_i(x) = W_i^T \phi_i(x) + \varepsilon_i(x), \quad i \in \mathbb{N}, \quad (18)$$

其中  $W_i \in \mathbb{R}^K$  为理想权值矢量,  $\phi_i \in \mathbb{R}^K$  为激活函数矢量,  $\varepsilon_i \in \mathbb{R}$  为重构误差,  $K$  为隐藏层神经元个数. 值函数关于  $x$  的微分为

$$\nabla V_i = \nabla \phi_i^T W_i + \nabla \varepsilon_i, \quad i \in \mathbb{N}, \quad (19)$$

其中  $\nabla \phi_i = \partial \phi_i / \partial x \in \mathbb{R}^{K \times 6}$ ,  $\nabla \varepsilon_i = \partial \varepsilon_i / \partial x \in \mathbb{R}^6$ . 根据式 (13), (14) 与 (19), 可得对应于理想权值矢量  $W_i$  的哈密顿函数为

$$\begin{aligned} H_i(x, W_1, \dots, W_N) &= x^T Q_i x + W_i^T \nabla \phi_i f(x) - \frac{1}{2} W_i^T \nabla \phi_i \sum_{j=1}^N g_j R_{jj}^{-1} g_j^T \nabla \phi_j^T W_j \\ &\quad + \frac{1}{4} \sum_{j=1}^N W_j^T \nabla \phi_j g_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} g_j^T \nabla \phi_j^T W_j \\ &\triangleq e_{Hi}, \quad i \in \mathbb{N}. \end{aligned} \quad (20)$$

由于理想权值矢量  $W_i$  事先未知, 因此记权值矢量估值为  $\hat{W}_i$ , 值函数的逼近值可表示为

$$\hat{V}_i(x) = \hat{W}_i^T \phi_i(x), \quad i \in \mathbb{N}. \quad (21)$$

值函数关于  $x$  的微分的逼近值为

$$\nabla \hat{V}_i = \nabla \phi_i^T \hat{W}_i, \quad i \in \mathbb{N}. \quad (22)$$

将式 (22) 代入 Nash 均衡策略的显式表达式 (14) 中, 可得

$$\hat{u}_i = -\frac{1}{2} R_{ii}^{-1} g_i^T \nabla \phi_i^T \hat{W}_i, \quad i \in \mathbb{N}. \quad (23)$$

哈密顿函数的逼近值为

$$\begin{aligned} H_i(x, \hat{W}_1, \dots, \hat{W}_N) &= x^T Q_i x + \hat{W}_i^T \nabla \phi_i f(x) - \frac{1}{2} \hat{W}_i^T \nabla \phi_i \sum_{j=1}^N g_j R_{jj}^{-1} g_j^T \nabla \phi_j^T \hat{W}_j \\ &\quad + \frac{1}{4} \sum_{j=1}^N \hat{W}_j^T \nabla \phi_j g_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} g_j^T \nabla \phi_j^T \hat{W}_j \\ &\triangleq e_i, \quad i \in \mathbb{N}. \end{aligned} \quad (24)$$

调整  $\hat{W}_i$  以最小化下述目标函数:

$$E_i = \frac{1}{2} e_i^2. \quad (25)$$

基于如下梯度下降法<sup>[15]</sup>更新评价神经网络权值矢量估值:

$$\dot{\hat{W}}_i = -\eta_i \frac{\partial E_i}{\partial \hat{W}_i} = -\eta_i e_i \theta_i, \quad (26)$$

其中  $\eta_i > 0$  为评价神经网络的学习速率,  $\theta_i = \nabla\phi_i(f(x) - \frac{1}{2}\sum_{j=1}^N g_j R_{jj}^{-1} g_j^T \nabla\phi_j^T \tilde{W}_j)$ . 评价神经网络权值矢量更新后, 可通过式 (23) 进行控制策略的更新, 因此可不引入执行神经网络, 以减小计算量.

式 (24) 中哈密尔顿函数的逼近表达式也为式 (15) 中 HJ 方程组的逼近表达式, 因此通过式 (26) 来调整神经网络权值矢量以最小化式 (25) 中目标函数的过程即为对 HJ 方程组的解逼近的过程. 在获得 HJ 方程组逼近解的基础上, 便可通过式 (23) 进一步获得各颗微小卫星的近似博弈均衡策略.

**定理1** 对于式 (4) 描述的系统, 评价神经网络权值矢量更新律由式 (26) 给出, 控制策略通过式 (23) 更新. 若由评价神经网络初始权值矢量能够得到初始可行控制策略集, 那么评价神经网络权值矢量估计误差能达到 UUB.

**证明** 选择如下 Lyapunov 函数:

$$L(t) = \sum_{i=1}^N L_i(t), \tag{27}$$

其中

$$L_i(t) = \frac{1}{2\eta_i} \tilde{W}_i^T \tilde{W}_i, \tag{28}$$

其导数为

$$\dot{L}_i(t) = \frac{1}{\eta_i} \tilde{W}_i^T \dot{\tilde{W}}_i, \tag{29}$$

且

$$\dot{\tilde{W}}_i = -\dot{W}_i = \eta_i e_i \theta_i, \tag{30}$$

因此,

$$\begin{aligned} \dot{L}_i(t) &= \tilde{W}_i^T \theta_i e_i \\ &= -(\tilde{W}_i^T \theta_i^*)(\tilde{W}_i^T \theta_i^*) - (\tilde{W}_i^T \theta_i^*) \left( \sum_{j=1}^N \tilde{W}_j^T D_{ij} \tilde{W}_j \right) - \left( \frac{1}{2} \sum_{j=1}^N \tilde{W}_i^T D_{ij} \tilde{W}_j \right) \left( \frac{1}{2} \sum_{j=1}^N \tilde{W}_i^T D_{ij} \tilde{W}_j \right) \\ &\quad + (\tilde{W}_i^T \theta_i) \left( \frac{1}{4} \sum_{j=1}^N \tilde{W}_j^T E_{ij} \tilde{W}_j \right) + (\tilde{W}_i^T \theta_i) \left( \frac{1}{2} \sum_{j=1}^N W_i^T D_{ij} \tilde{W}_j \right) \\ &\quad - (\tilde{W}_i^T \theta_i) \left( \frac{1}{2} \sum_{j=1}^N W_j^T E_{ij} \tilde{W}_j \right) + (\tilde{W}_i^T \theta_i) e_{Hi}, \end{aligned} \tag{31}$$

其中

$$\begin{aligned} D_{ij} &= \nabla\phi_i g_j R_{jj}^{-1} g_j^T \nabla\phi_j^T, \\ E_{ij} &= \nabla\phi_j g_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} g_j^T \nabla\phi_j^T. \end{aligned} \tag{32}$$

假设  $\|D_{ij}\| \leq \lambda_{DMij}$ ,  $\|E_{ij}\| \leq \lambda_{EMij}$ ,  $\|e_{Hi}\| \leq b_{eHi}$ ,  $\|W_i^T D_{ij}\| \leq b_{Dij}$ ,  $\|W_j^T E_{ij}\| \leq b_{Eij}$ ,  $\theta_{im} \leq \|\theta_i^*\| \leq \theta_{iM}$ ,  $b_{im} \leq \|\theta_i\| \leq b_{iM}$ , 其中  $\lambda_{DMij}$ ,  $\lambda_{EMij}$ ,  $b_{eHi}$ ,  $b_{Dij}$ ,  $b_{Eij}$ ,  $\theta_{im}$ ,  $\theta_{iM}$ ,  $b_{im}$ ,  $b_{iM}$  均为正常数. 其中, 假设  $\theta_{im} \leq \|\theta_i^*\|$  意味着系统状态需要满足持续激励条件.



考虑如下的关系式:

$$ab = \frac{1}{2} \left( - \left( \psi_+ a - \frac{b}{\psi_+} \right)^2 + \psi_+^2 a^2 + \frac{b^2}{\psi_+^2} \right), \quad (33a)$$

$$-ab = -\frac{1}{2} \left( \left( \psi_- a + \frac{b}{\psi_-} \right)^2 - \psi_-^2 a^2 - \frac{b^2}{\psi_-^2} \right), \quad (33b)$$

其中  $\psi_+$  与  $\psi_-$  为可调系数, 均为非零常数.

对于式 (31) 中的第 2 项来说,

$$-(\tilde{W}_i^T \theta_i^*) \left( \sum_{j=1}^N \tilde{W}_i^T D_{ij} \tilde{W}_j \right) = -(\tilde{W}_i^T \theta_i^*) (\tilde{W}_i^T D_{i1} \tilde{W}_1 + \dots + \tilde{W}_i^T D_{iN} \tilde{W}_N). \quad (34)$$

根据式 (33) 中的关系式, 可得

$$\begin{aligned} -(\tilde{W}_i^T \theta_i^*) (\tilde{W}_i^T D_{ik_1} \tilde{W}_{k_1}) &= -\frac{1}{2} \left( \left( \psi_{i2k_1} (\tilde{W}_i^T \theta_i^*) + \frac{1}{\psi_{i2k_1}} (\tilde{W}_i^T D_{ik_1} \tilde{W}_{k_1}) \right)^2 \right. \\ &\quad \left. - \psi_{i2k_1}^2 (\tilde{W}_i^T \theta_i^*)^2 - \frac{1}{\psi_{i2k_1}^2} (\tilde{W}_i^T D_{ik_1} \tilde{W}_{k_1})^2 \right) \\ &\leq \frac{\psi_{i2k_1}^2 \theta_{iM}^2}{2} \|\tilde{W}_i\|^2 + \frac{\lambda_{DMik_1}^2}{2\psi_{i2k_1}^2} \|\tilde{W}_i\|^2 \|\tilde{W}_{k_1}\|^2, \end{aligned} \quad (35)$$

$\psi_{i2k_1}$  对应于式 (33) 中的可调系数. 则有

$$-(\tilde{W}_i^T \theta_i^*) \left( \sum_{j=1}^N \tilde{W}_i^T D_{ij} \tilde{W}_j \right) \leq \sum_{j=1}^N \frac{\psi_{i2j}^2 \theta_{iM}^2}{2} \|\tilde{W}_i\|^2 + \sum_{j=1}^N \frac{\lambda_{DMij}^2}{2\psi_{i2j}^2} \|\tilde{W}_i\|^2 \|\tilde{W}_j\|^2. \quad (36)$$

对第 4~7 项可作类似的处理. 对第 1 项作如下处理:

$$-(\tilde{W}_i^T \theta_i^*) (\tilde{W}_i^T \theta_i^*) \leq -\theta_{im}^2 \|\tilde{W}_i\|^2. \quad (37)$$

对第 3 项来说, 由于

$$-\left( \frac{1}{2} \sum_{j=1}^N \tilde{W}_i^T D_{ij} \tilde{W}_j \right) \left( \frac{1}{2} \sum_{j=1}^N \tilde{W}_i^T D_{ij} \tilde{W}_j \right) < 0,$$

因此存在正定矩阵  $M_1$ , 使得

$$\sum_{i=1}^N \left( - \left( \frac{1}{2} \sum_{j=1}^N \tilde{W}_i^T D_{ij} \tilde{W}_j \right) \left( \frac{1}{2} \sum_{j=1}^N \tilde{W}_i^T D_{ij} \tilde{W}_j \right) \right) \leq -Z^T M_1 Z, \quad (38)$$

其中  $Z \triangleq [\|\tilde{W}_1\|^2 \dots \|\tilde{W}_N\|^2]^T \in \mathbb{R}^N$ .

对式 (31) 中的各项进行上述处理后可得

$$\dot{L}(t) = \sum_{i=1}^N \dot{L}_i(t) \leq -Z^T M Z + T Z + K, \quad (39)$$

其中

$$\begin{aligned} M &= M_1 - \Psi_{21} - \Psi_{41}, \\ T &= \Psi_1 + \Psi_{22} + \Psi_{42} + \Psi_5 + \Psi_6 + \Psi_{71}, \\ K &= \Psi_{72}, \end{aligned}$$

$\Psi_{(\cdot)}$  的表达式见附录 A. 由于通过调整  $\Psi_{21}$ ,  $\Psi_{41}$  中的可调系数  $\psi_{(\cdot)}$ ,  $M$  的正定性可得到保证, 因此, 一旦不等式

$$\|Z\| > \frac{T_M + \sqrt{T_M^2 + 4\lambda_M K_M}}{2\lambda_m} \quad (40)$$

成立,  $\dot{L} < 0$ . 其中  $\lambda_m \leq \|M\| \leq \lambda_M$ ,  $\|T\| \leq T_M$ ,  $\|K\| \leq K_M$ . 由 Lyapunov 定理可知,  $\|Z\|$  是 UUB 的, 进而可知  $\|\tilde{W}_i\|$  也是 UUB 的.

**定理 2** 对于式 (4) 描述的系统, 评价神经网络权值矢量更新律由式 (26) 给出, 控制策略通过式 (23) 更新. 若由评价神经网络初始权值矢量能够得到初始可行控制策略集, 那么对于初始时刻  $t_0$  和初始条件  $x_0$ , 存在时间  $T(x_0, B)$ , 使得  $x(t)$  是 UUB 的. 即

$$\begin{aligned} \|x(t)\| &\leq \max \left\{ \sqrt{\frac{\xi_1}{\lambda_{\min}\{Q_1\}}}, \dots, \sqrt{\frac{\xi_N}{\lambda_{\min}\{Q_N\}}} \right\} \\ &\triangleq B, \quad t \geq t_0 + T, \end{aligned} \quad (41)$$

其中  $\xi_i > 0, i \in \mathbb{N}$ , 定义如下:

$$\left\| \frac{1}{2} (\nabla \phi_i^T W_i + \nabla \varepsilon_i)^T \sum_{j=1}^N g_j R_{jj}^{-1} g_j^T (\nabla \phi_j^T \tilde{W}_j + \nabla \varepsilon_j) \right\| \leq \xi_i. \quad (42)$$

定理 2 的证明可参考文献 [15].

定理 2 意味着组合体姿态状态量最后会收敛到靠近零值的邻域内, 且该邻域大小与正数  $\xi_i$  和矩阵  $Q_i$  最小特征值有关.  $\xi_i$  为式 (42) 中不等号左端函数范数的上限, 由于定理 1 中已经证明  $\tilde{W}_i$  是 UUB 的, 因此, 式 (42) 中不等号左端函数范数是有界的,  $\xi_i$  也是有界的. 由于矩阵  $Q_i$  是可选择的, 通过选择具有尽可能大的最小特征值的矩阵  $Q_i$ , 可使得组合体姿态状态量尽可能好地趋近于零, 因此, 可通过选择合适的  $Q_i$  以获得期望的控制精度.

## 5 仿真

本节通过数值仿真对失效航天器姿态接管的微分博弈控制方法的有效性进行验证. 仿真分为两阶段, 第 1 阶段为评价神经网络权值矢量学习阶段, 第 2 阶段为姿态接管控制阶段.

### 5.1 权值矢量学习阶段

假设有三颗微小卫星进行失效航天器的姿态接管控制, 其各自本体坐标系到参考坐标系的转换矩阵分别为

$$C_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 0.8829 & 0 & 0.4695 \\ 0.4695 & 0 & -0.8829 \\ 0 & 1 & 0 \end{bmatrix}, \quad C_3 = \begin{bmatrix} 0.7986 & -0.6018 & 0 \\ -0.6018 & -0.7986 & 0 \\ 0 & 0 & -1 \end{bmatrix},$$

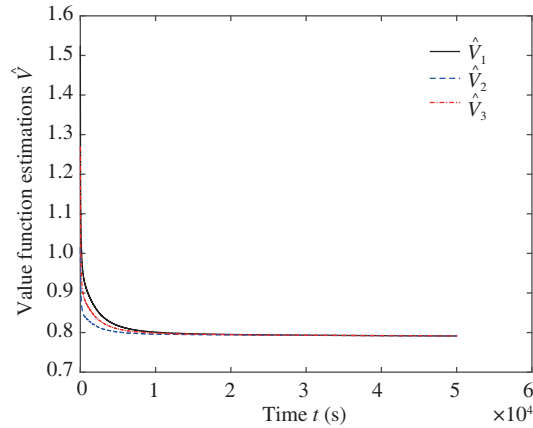


图 2 (网络版彩图) 微小卫星值函数变化曲线

Figure 2 (Color online) Value function trajectories of microsattellites

组合体转动惯量矩阵为

$$J = \begin{bmatrix} 90 & 10 & 5 \\ 10 & 50 & 7 \\ 5 & 7 & 50 \end{bmatrix},$$

各微小卫星评价神经网络的基函数为

$$\phi_i = \left[ \frac{1}{2}x_1^2 \frac{1}{2}x_2^2 \frac{1}{2}x_3^2 \frac{1}{2}x_4^2 \frac{1}{2}x_5^2 \frac{1}{2}x_6^2 x_1x_4 x_2x_5 x_3x_6 \right]^T, \quad i \in \mathbb{N}.$$

选择微小卫星性能指标函数中的矩阵为  $Q_1 = I_6$ ,  $Q_2 = I_6$ ,  $Q_3 = I_6$ ,  $R_{11} = R_{12} = R_{13} = 0.01I_3$ ,  $R_{21} = R_{22} = R_{23} = 0.01I_3$ ,  $R_{31} = R_{32} = R_{33} = 0.01I_3$ . 三颗微小卫星评价神经网络初始权值矢量估值分别选择为  $\hat{W}_1^0 = 3019$ ,  $\hat{W}_2^0 = 2019$ ,  $\hat{W}_3^0 = 2519$ . 为了保证持续激励条件得到满足, 在权值矢量学习阶段, 每隔 10 s 将组合体的状态量重置为  $x_{PE} = [0.0324, 0.0373, 0.0324, 0.15, 0.15, 0.15]^T$ .

图 2 为评价神经网络权值矢量收敛过程中三颗微小卫星在组合体状态量为  $x = x_{PE}$  时的值函数的变化曲线, 可以看出, 随着神经网络权值矢量估值的收敛, 各微小卫星的值函数逐渐得到优化并减小. 图 3 为三颗微小卫星评价神经网络权值矢量估值的收敛曲线. 从图中可以看出, 权值矢量的学习需要较长时间, 但一旦学习完成后, 微小卫星的控制策略即可通过式 (23) 直接计算得到, 因此之后将能够在不需要进行控制分配的情况下实施失效航天器的姿态接管控制.

## 5.2 姿态接管控制阶段

考虑到微小卫星控制能力的有限性, 本文考虑小角度的姿态镇定问题. 假设在姿态接管控制阶段, 组合体的状态量初值为  $x_0 = [0.0015, 0.0009, 0.0013, 0, 0, 0]^T$ . 图 4 和 5 分别为失效航天器姿态接管控制阶段组合体姿态 MRPs 和姿态角速度的变化曲线, 可以看出基于学习到的权值矢量, 微小卫星实现了失效航天器的姿态镇定.

姿态接管控制阶段三颗微小卫星控制力矩的变化曲线如图 6 所示. 从图中可以看出, 在初始控制阶段, 由于状态量幅值较大, 因此所需的控制力矩较大, 随着状态量幅值的不断减小, 所需的控制力矩也逐渐减小.

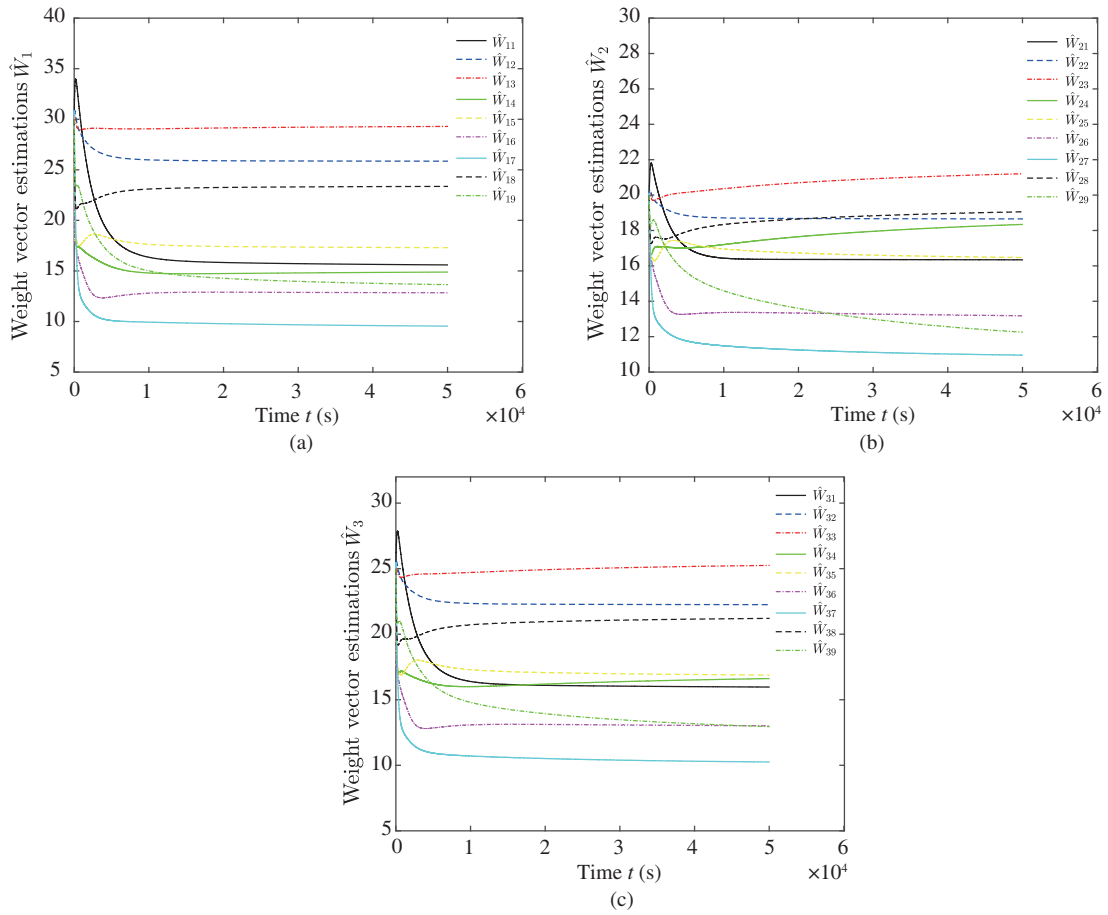


图 3 (网络版彩图) 微小卫星评价神经网络权值矢量估值收敛曲线

Figure 3 (Color online) Convergence of critic NN weight vector estimations of microsatellites. (a) Microsatellite 1; (b) Microsatellite 2; (c) Microsatellite 3

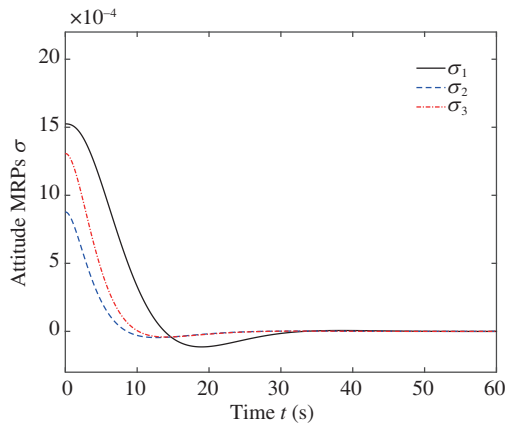


图 4 (网络版彩图) 组合体姿态 MRPs 变化曲线

Figure 4 (Color online) Attitude MRPs trajectories of combination

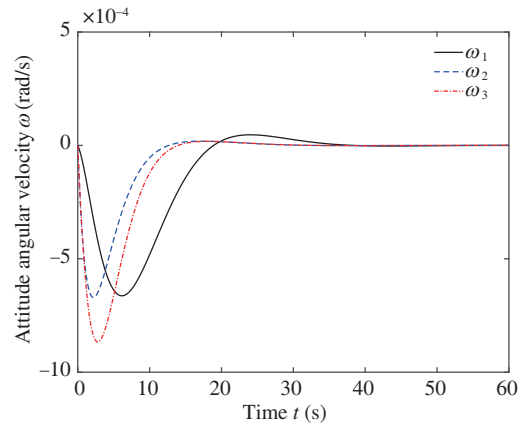


图 5 (网络版彩图) 组合体姿态角速度变化曲线

Figure 5 (Color online) Attitude angular velocity trajectories of combination

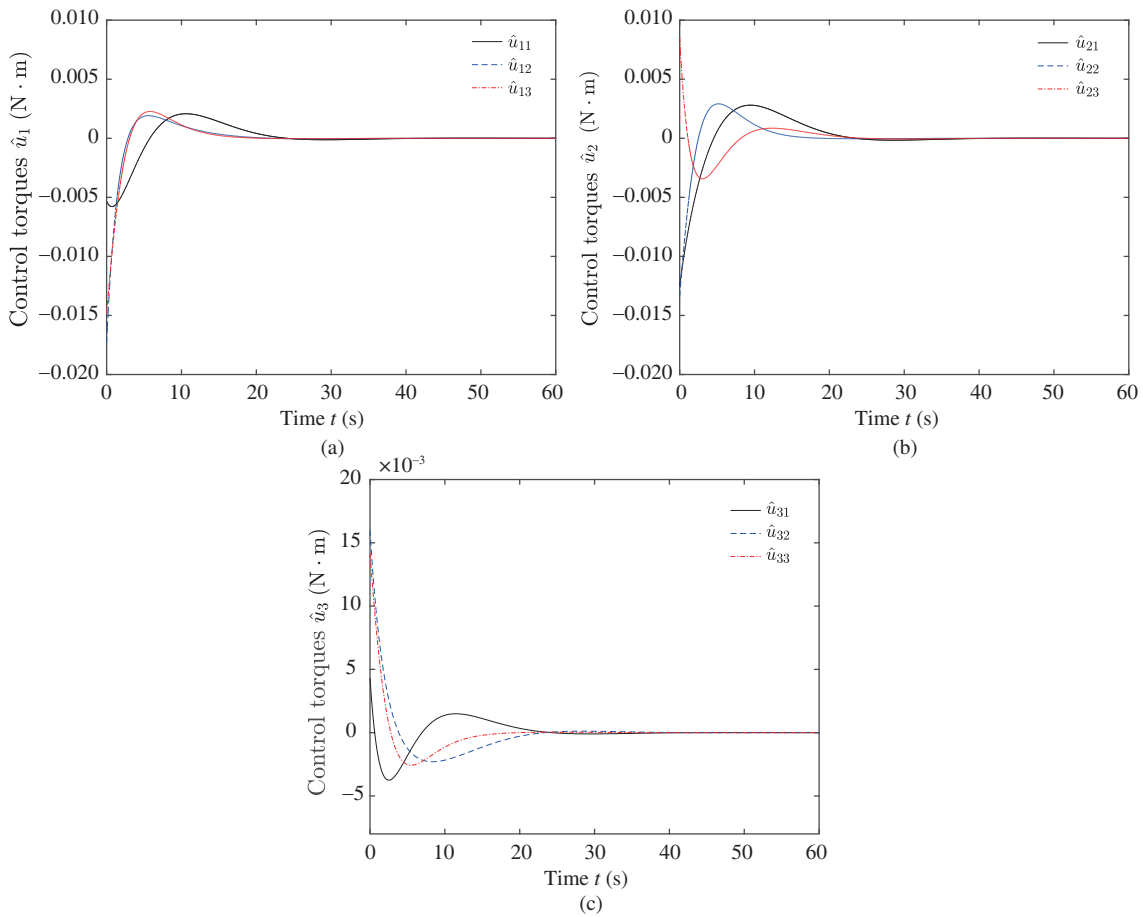


图 6 (网络版彩图) 微小卫星控制力矩变化曲线

Figure 6 (Color online) Control torques trajectories of microsatellites. (a) Microsatellite 1; (b) Microsatellite 2; (c) Microsatellite 3

## 6 结论

本文通过微分博弈方法研究了失效航天器的姿态接管控制问题. 根据组合体的姿态运动学模型和姿态动力学模型建立了多微小卫星的微分博弈模型, 通过基于单神经网络的策略迭代算法对微小卫星的博弈均衡策略进行了逼近, 通过 Lyapunov 方法证明了评价神经网络权值矢量估计误差的一致最终有界性. 在数值仿真中, 基于学习到的权值矢量, 在不需要进行控制分配的情况下通过微小卫星的自主决策实现了对失效航天器的姿态接管控制. 本文的工作是对基于微分博弈方法实现微小卫星自主协同决策的尝试, 未来的研究工作中, 我们将进一步考虑怎样在具有干扰及不确定性的环境中通过微小卫星的博弈决策实现如在轨协同装配、协同运输等更加复杂的空间协同操控任务.

## 参考文献

- 1 Zhai G, Zhang J R, Zhou Z C. A review of on-orbit life-time extension technologies for GEO satellites. *J Astronaut*, 2012, 33: 849–859
- 2 Huang P F, Wang M, Chang H T, et al. Takeover control of attitude maneuver for failed spacecraft. *J Astronaut*, 2016, 37: 924–935

- 3 Wang Z, Yuan J, Shi Y, et al. Robust adaptive fault tolerant attitude control for post-capture non-cooperative targets with actuator nonlinearities. *Trans Inst Meas Control*, 2018, 40: 2116–2128
- 4 Chang H T, Huang P F, Wang M, et al. Distributed control allocation for cellular space robots in takeover control. *Acta Aeronaut Astronaut Sin*, 2016, 37: 2864–2873 [常海涛, 黄攀峰, 王明, 等. 空间细胞机器人接管控制的分布式控制分配. *航空学报*, 2016, 37: 2864–2873]
- 5 Jaeger T, Mirczak W. Satlets - the building blocks of future satellites - and which mold do you use? In: *Proceedings of AIAA SPACE 2013 Conference and Exposition*, San Diego, 2013
- 6 Goeller M, Oberlaender J, Uhl K, et al. Modular robots for on-orbit satellite servicing. In: *Proceedings of IEEE International Conference on Robotics and Biomimetics*, Guangzhou, 2012. 2018–2023
- 7 Xue L, Wang Q L, Sun C Y. Game theoretical approach for the leader selection of the second-order multi-agent system. *Control Theory Appl*, 2016, 33: 1593–1602 [薛磊, 王庆领, 孙长银. 博弈论框架下的二阶多智能体系统领导者选择算法. *控制理论与应用*, 2016, 33: 1593–1602]
- 8 Lin W. Distributed UAV formation control using differential game approach. *Aerosp Sci Tech*, 2014, 35: 54–62
- 9 Vamvoudakis K G, Lewis F L, Hudus G R. Multi-agent differential graphical games: online adaptive learning solution for synchronization with optimality. *Automatica*, 2012, 48: 1598–1611
- 10 Abouheaf M I, Lewis F L, Vamvoudakis K G, et al. Multi-agent discrete-time graphical games and reinforcement learning solutions. *Automatica*, 2014, 50: 3038–3053
- 11 Ru C J, Wei R X, Guo Q, et al. Guidance control of cognitive game for unmanned aerial vehicle autonomous collision avoidance. *Control Theor Appl*, 2014, 31: 1555–1560 [茹常剑, 魏瑞轩, 郭庆, 等. 面向无人机自主防碰撞的认知博弈制导控制. *控制理论与应用*, 2014, 31: 1555–1560]
- 12 Bopardikar S D, Bullo F, Hespanha J P. On discrete-time pursuit-evasion games with sensing limitations. *IEEE Trans Robot*, 2008, 24: 1429–1439
- 13 Blasch E P, Pham K, Shen D. Orbital satellite pursuit-evasion game-theoretical control. In: *Proceedings of the 11th International Conference on Information Science, Signal Processing and Their Applications*, Montreal, 2012. 1007–1012
- 14 Vamvoudakis K G, Lewis F L. Multi-player non-zero-sum games: online adaptive learning solution of coupled Hamilton-Jacobi equations. *Automatica*, 2011, 47: 1556–1569
- 15 Liu D, Li H, Wang D. Online synchronous approximate optimal learning algorithm for multi-player non-zero-sum games with unknown dynamics. *IEEE Trans Syst Man Cybern Syst*, 2014, 44: 1015–1027
- 16 Vamvoudakis K G. Non-zero sum Nash Q-learning for unknown deterministic continuous-time linear systems. *Automatica*, 2015, 61: 274–281
- 17 Schaub H, Junkins J L. Analytical mechanics of space systems. In: *Proceedings of AIAA Education Series*, Reston, 2003. 107–142

## 附录 A 可调系数

$$\Psi_1 = [-\theta_{1m}^2 \cdots -\theta_{Nm}^2], \quad \Psi_{21} = \begin{bmatrix} \frac{\lambda_{DM11}^2}{2\psi_{121}^2} & \cdots & \frac{\lambda_{DM1N}^2}{2\psi_{12N}^2} \\ \vdots & & \vdots \\ \frac{\lambda_{DMN1}^2}{2\psi_{N21}^2} & \cdots & \frac{\lambda_{DMNN}^2}{2\psi_{N2N}^2} \end{bmatrix}, \quad \Psi_{22} = \left[ \sum_{j=1}^N \frac{\psi_{12j}^2 \theta_{1M}^2}{2} \cdots \sum_{j=1}^N \frac{\psi_{N2j}^2 \theta_{NM}^2}{2} \right],$$

$$\Psi_{41} = \begin{bmatrix} \sum_{j=1}^N \frac{\lambda_{EMj1}^2}{32\psi_{j41}^2} & & \\ & \ddots & \\ & & \sum_{j=1}^N \frac{\lambda_{EMjN}^2}{32\psi_{j4N}^2} \end{bmatrix}, \quad \Psi_{42} = \left[ \sum_{j=1}^N \frac{\psi_{14j}^2 b_{1M}^2}{2} \cdots \sum_{j=1}^N \frac{\psi_{N4j}^2 b_{NM}^2}{2} \right],$$

$$\Psi_5 = \left[ \sum_{j=1}^N \frac{\psi_{15j}^2 b_{1M}^2}{2} \cdots \sum_{j=1}^N \frac{\psi_{N5j}^2 b_{NM}^2}{2} \right] + \left[ \sum_{j=1}^N \frac{b_{Dj1}^2}{8\psi_{j51}^2} \cdots \sum_{j=1}^N \frac{b_{DjN}^2}{8\psi_{j5N}^2} \right],$$

$$\Psi_6 = \left[ \sum_{j=1}^N \frac{\psi_{16j}^2 b_{1M}^2}{2} \cdots \sum_{j=1}^N \frac{\psi_{N6j}^2 b_{NM}^2}{2} \right] + \left[ \sum_{j=1}^N \frac{b_{Ej1}^2}{8\psi_{j61}^2} \cdots \sum_{j=1}^N \frac{b_{EjN}^2}{8\psi_{j6N}^2} \right],$$

$$\Psi_{71} = \left[ \frac{\psi_{17}^2 b_{1M}^2}{2} \dots \frac{\psi_{N7}^2 b_{NM}^2}{2} \right], \quad \Psi_{72} = \sum_{i=1}^N \frac{b_{e_{Hi}}^2}{2\psi_{i7}^2},$$

其中  $\Psi_{21}, \Psi_{41} \in \mathbb{R}^{N \times N}$ ,  $\Psi_1, \Psi_{22}, \Psi_{42}, \Psi_5, \Psi_6, \Psi_{71} \in \mathbb{R}^{1 \times N}$ ,  $\Psi_{72} \in \mathbb{R}$ .

## Differential game learning approach for multiple microsatellites takeover of the attitude movement of failed spacecraft

Nan HAN<sup>1,2</sup>, Jianjun LUO<sup>1,2\*</sup> & Yuan CHAI<sup>1,2</sup>

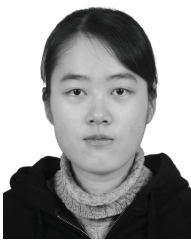
1. *Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen 518057, China;*

2. *Science and Technology on Aerospace Flight Dynamics Laboratory, School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China*

\* Corresponding author. E-mail: jjluo@nwpu.edu.cn

**Abstract** Takeover of the attitude control function of a failed spacecraft suffering from fuel exhaustion or actuator failures enables recycling of on-board valuable reusable payloads. Microsatellites can provide cost-efficient ways for the attitude takeover control through coordination. Differential games are used to study the individual optimal decision problem, where each player optimizes their local performance index function to obtain the control policy, and the game's predefined global objective can be achieved. In this paper, the failed spacecraft attitude takeover control problem is transformed into a multi-microsatellite differential game problem. First, the multi-microsatellite differential game model is established, the performance index function is designed for each microsatellite, and the mathematical description of the multi-microsatellite differential game problem is realized. Second, the Hamilton-Jacobi (HJ) equations are provided and solved through the single neural network (NN) based policy iteration (PI) algorithm to learn the multi-microsatellite game equilibrium control strategies. Finally, numerical simulations are carried out to validate the effectiveness of the multi-microsatellite differential game learning method. The results have shown that the predefined global objective of the takeover of the attitude control of the failed spacecraft can be realized through the approximate game equilibrium control strategies of multiple microsatellites.

**Keywords** attitude control, takeover control, differential games, game control, policy iteration



**Nan HAN** was born in 1992. She received an M.S. degree from the Northwestern Polytechnical University, Xi'an, in 2016. Currently, she is a Ph.D. candidate at the Northwestern Polytechnical University. Her research interests include distributed systems and multi-agent coordinated control.



**Jianjun LUO** was born in 1965. He received a Ph.D. degree from Northwestern Polytechnical University, Xi'an, in 1999. Currently, he is a professor at the Northwestern Polytechnical University. His research interests include aerospace flight dynamics, navigation, and guidance.