

УДК 629.785,004.032.26

НЕЙРОАДАПТИВНОЕ ПОДДЕРЖАНИЕ ФОРМАЦИИ СПУТНИКОВ НА НИЗКИХ ОКОЛОЗЕМНЫХ ОРБИТАХ

© 2022 г. М. Г. Ширококов¹, *, С. П. Трофимов¹

¹Институт прикладной математики им. М.В. Келдыша РАН, Москва, Россия

*shirobokov@keldysh.ru

Поступила в редакцию 27.11.2020 г.

После доработки 12.02.2021 г.

Принята к публикации 16.06.2021 г.

Работа посвящена разработке и исследованию нейроадаптивного управления формацией из двух спутников на низкой околоземной орбите. Цель управления – поддержание проективной круговой орбиты, при этом считается, что один из аппаратов неуправляемый и его баллистический коэффициент не известен, а другой аппарат – управляемый, при этом есть возможность управлять как его ориентацией (а значит, величиной миделева сечения), так и орбитальным движением с помощью двигателя большой тяги. Искусственные нейронные сети используются для аппроксимации построенных функций управления, а адаптивность системы управления выражается в адаптации управления к неизвестному баллистическому коэффициенту неуправляемого аппарата. В работе детально описывается реализация, архитектура и процесс обучения нейронных сетей, а также способ оценки их качества. Дается описание процедуры адаптации к неизвестному параметру неуправляемого аппарата в режиме реального времени. Показано, что адаптация возможна в течение приемлемого времени. Также приводятся результаты адаптации к двум параметрам модели – баллистическому коэффициенту неуправляемого аппарата и плотности атмосферы. Моделирование производится в рамках возмущенной задачи двух тел с учетом силы атмосферного сопротивления и влияния от второй зональной гармоники геопотенциала.

DOI: 10.31857/S0023420622010101

1. ВВЕДЕНИЕ

В настоящее время методы машинного обучения активно применяются для решения самых разнообразных задач человеческой деятельности. В первую очередь это относится к задачам распознавания образов, прогнозирования эволюции динамических систем и синтеза управления ими, системам искусственного интеллекта, теории принятия решений. Методы машинного обучения сейчас можно обнаружить в экономике, медицине, биохимических исследованиях, системах фильтрации информации и др.

Методы машинного обучения принято делить на три типа: методы обучения с учителем, методы обучения без учителя и методы обучения с подкреплением [1, 2]. Методы обучения с учителем применяются в случаях, когда требуется аппроксимировать зависимость между какими-либо переменными задачи. В этом случае составляется параметрическая модель этой зависимости и множество конкретных пар этих переменных, которое используется для подгонки параметров модели. Процесс такой подгонки называется *обучением модели*, а множество пар переменных – *обучающей выборкой*. Методы обучения без учителя

используются в случаях, когда дано некоторое множество значений какой-то одной переменной и необходимо определить структуру этого множества (разбиение на подгруппы, принадлежность подпространству меньшей размерности и пр.). Методы обучения с подкреплением созданы для ситуаций, когда исследуемая система может быть формализована как состоящая из *среды* и *агента*, действующего в данной среде и получающего от среды скалярный сигнал, который он пытается методом проб и ошибок максимизировать. Поведение агента в среде определяется его стратегией – отображением из множества состояний среды в множество действий. Задача обучения с подкреплением состоит в том, чтобы определить оптимальную стратегию, т.е. стратегию, которая максимизирует суммарный получаемый сигнал подкрепления от среды. Стратегия, как отображение, обычно аппроксимируется какой-либо параметрической моделью.

Уже несколько лет в литературе наблюдается тренд на применение методов машинного обучения для проектирования схем управления космическими аппаратами. Эти работы можно условно разделить на несколько классов: 1) синтез конту-

ра управления орбитальным и/или угловым движением космического аппарата, 2) поиск оптимальных управляющих воздействий или оптимальных траекторий перелета на этапе предварительного анализа миссии, 3) распознавание образов для решения задач навигации или дистанционного зондирования, 4) диагностика аппаратуры и 5) все остальные работы, которые трудно классифицировать.

Данная работа относится к первому из перечисленных классов и ставит своей целью оценку возможностей управления формацией спутников на низких околоземных орбитах с помощью искусственных нейронных сетей (далее просто нейронных сетей), а также оценку возможностей адаптации системы управления к меняющимся параметрам аппаратов и среды в режиме полета. Искусственные нейронные сети и выступают, как правило, в роли модели, призванной аппроксимировать зависимость между изучаемыми переменными в методах обучения с учителем или для аппроксимации оптимальной стратегии в методах обучения с подкреплением.

Исследования, в которых методы машинного обучения используются для проектирования оптимальных траекторий или синтеза контура управления, широко представлены в современной литературе. Их можно разбить на два класса: те, что используют методы обучения с учителем, и те, что проводятся с помощью методов обучения с подкреплением. Методы обучения без учителя в задачах управления космическим аппаратом авторам не встречались.

Отметим только ключевые (на взгляд авторов) исследования. Во-первых, упомянем работы [3, 4], в которых строится адаптивное управление вращательным движением космического аппарата с использованием нейронных сетей. Идея состоит в том, чтобы ввести две нейронных сети – прогнозирующую и управляющую. Управляющая нейронная сеть ответственна за выработку управляющих воздействий на аппарат на основе входа – состояния аппарата. Прогнозирующая сеть помогает понять, насколько выбранные управляющие воздействия в текущем состоянии оказываются эффективными: делается это для того, чтобы на основании оценок эффективности скорректировать управляющую нейронную сеть для лучшей работы. Прогнозирующая и управляющая нейронные сети вводятся и в работе [5], где с использованием нейронных сетей решается задача адаптации параметров системы управления реального космического аппарата *Solar and Heliospheric Observatory (SOHO)* [6].

Надо сказать, что идея использования прогнозирующей сети для коррекции управляющей сети критикуется уже в одной из первых работ [3]. Проблема подобных коррекций связана с тем, что

прогнозирующая нейронная сеть должна быть достаточно точной и сама должна корректироваться чаще, чем управляющая сеть. Кроме того, коррекция весов управляющей сети во время полета способна существенно дестабилизировать управление. Поэтому, несмотря на то что в предыдущих наших работах [7, 8] решение задач управления одиночными аппаратами с большой и малой тягой было произведено с помощью управляющей и прогнозирующей нейронных сетей, многочисленные численные эксперименты убедили нас отказаться от этого способа адаптации нейронных сетей в текущей работе.

В работах [9, 10] решается задача аппроксимации функции оптимального управления малой тягой в орбитальном полете. Идея данных работ состоит в замыкании в единый контур нейронных сетей, выдающих значения сопряженных переменных, и обычного метода оптимизации, решающего краевую задачу из принципа максимума Понтрягина. Выход нейронных сетей используется как начальное приближение для метода решения краевой задачи, а выход обычного оптимизационного метода – как правильный пример для коррекции нейронных сетей. Кроме того, интересна идея выделения под каждый аппроксимируемый параметр своей нейронной сети. Этим приемом мы воспользуемся и в данной работе.

В работе [11] решается задача управления формацией с субмиллиметровой точностью с использованием методов детерминированного обучения – т.е. методов обучения с учителем, регулятор которых синтезируется с помощью методов Ляпунова и на основе нейронных сетей, веса которых подчиняются некоторым дифференциальным уравнениям, гарантирующим сходимость к оптимальным (в том или ином смысле) значениям. При этом регулятор на основе нейронных сетей является здесь помощником для основного регулятора (линейно-квадратичный регулятор) и способен подавить возмущения от сил светового давления и гравитации Луны. Основным регулятором в методах детерминированного обучения могут служить и другие регуляторы, например на основе скользящего управления [12].

Все перечисленные выше работы используют методы обучения с учителем для обучения моделей, аппроксимирующих оптимальные управляющие воздействия на космический аппарат. Что касается методов обучения с подкреплением, имеется большая группа работ немецкого специалиста в области прикладной математики Б. Давхальда и его коллег [13–16]. Все эти работы нацелены на поиск оптимальных траекторий перелета в различных постановках задач. Основная идея состоит в том, чтобы аппроксимировать стратегию поведения агента (функцию управления) с помощью нейронных сетей, а их веса искать гене-

тическими методами оптимизации. Эта связь между машинным обучением и эволюционным программированием особенно подчеркивается и обсуждается в обзорной работе [17]. Похожий метод также описан в работе [18], в которой ставится задача прецизионного управления формацией спутников, где для оптимизации применяется не генетический алгоритм, а метод роя частиц. Предлагаемый подход был протестирован на платформе Synchronized Position Hold Engage Reorient Experimental Satellites (SPHERES) [19] на борту *Международной космической станции*.

Наконец, отметим еще один крупный класс работ, использующих методы обучения с подкреплением. Это работы американских специалистов Р. Фурфаро и Б. Гаудета [20–24]. Все они посвящены задаче адаптивного управления аппаратом во время спуска на поверхность Марса или другого небесного тела. Оптимальное управление в первых двух работах ищется прямым методом (параметризация функции управления и оптимизация ее параметров). В трех последних работах оптимальное управление ищется непрямыми методами, т.е. таким образом, чтобы оно удовлетворяло уравнению оптимальности Беллмана.

Методы обучения с подкреплением, особенно в задачах с непрерывным множеством состояний и управляющих воздействий, характеризуются повышенной сложностью реализации подобных систем обучения и значительные затраты времени на обучение нейронных сетей даже в сравнительно простых случаях – например, задаче удержания математического маятника в неустойчивом положении равновесия. Многочисленные попытки авторов решить задачу текущей работы методами обучения с подкреплением не увенчались успехом, но результативными оказались методы обучения с учителем.

Итак, в работе разрабатывается адаптивная система управления формацией спутников, движущихся на низких околокруговых орбитах вокруг Земли, на основе нейронных сетей и с использованием силы атмосферного сопротивления. В разделе 2 описывается модель, в рамках которой рассматривается движение космических аппаратов. В разделе 3 ставится задача поддержания спутниковой формации, предлагается метод ее решения, описывается возникающая при этом оптимизационная процедура. В результате решения оптимизационной задачи для различных начальных данных формируется выборка, на которой обучаются две предлагаемые нами управляющие нейронные сети. Способ формирования выборки, метод обучения, архитектура нейронных сетей и результаты расчетов приводятся в разделе 4. В разделе 5 описывается метод адаптации системы управления к меняющимся параметрам спутников и плотности атмосферы, при-

водятся результаты работы метода. Далее следует заключение, в котором собраны все основные выводы исследования.

2. ОСНОВНЫЕ ОБОЗНАЧЕНИЯ И МОДЕЛИ ДВИЖЕНИЯ

Рассматривается движение двух спутников вблизи круговой орбиты радиуса R_{ref} вокруг Земли. Орбитальную скорость движения по орбите с радиусом R_{ref} обозначим $V_{\text{ref}} = \sqrt{\mu_E/R_{\text{ref}}}$, где μ_E – гравитационный параметр Земли. Будем считать, что на спутники действует только сила гравитационного притяжения Земли и сила атмосферного сопротивления, а движение спутников определяется дифференциальными уравнениями

$$\dot{\mathbf{R}}_i = \mathbf{V}_i, \quad \mathbf{V}_i = -\frac{\mu_E}{R_i^3} \mathbf{R}_i + \mathbf{F}_{a,i} + \mathbf{F}_{J_{2,i}}, \quad i = 1, 2, \quad (1)$$

записанными в произвольной инерциальной системе координат $CXYZ$ с началом в центре масс Земли; для определенности будем считать, что плоскость CXY инерциальной системы координат совпадает с плоскостью экватора Земли, а ось CZ направлена вдоль оси вращения Земли. Здесь $\mathbf{R}_i = [X_i, Y_i, Z_i]$ – радиус-вектор i -го спутника, \mathbf{V}_i – вектор его скорости, $\mathbf{F}_{a,i}$ – ускорение вследствие торможения в атмосфере, определяемое согласно кусочно-экспоненциальной модели атмосферы и вычисляемое по формуле

$$\mathbf{F}_{a,i} = -\frac{1}{2} \cdot \beta_i \cdot \rho \cdot V_i^2 \cdot \frac{\mathbf{V}_i}{V_i}, \quad \rho = \rho_l \left(\frac{\rho_h}{\rho_l} \right)^{(h_i - h_0)/100},$$

где $\beta_i = c_{x,i}(A_i/m_i)$ – баллистический коэффициент i -го спутника, т.е. произведение коэффициента аэродинамического сопротивления $c_{x,i}$ и отношения площади A_i к массе m_i , а параметры ρ_l , ρ_h и h_0 зависят от высоты спутника $h_i = R_i - R_E$, R_E – радиус Земли; параметр h_0 и высота h_i в данной формуле выражены в километрах. Зависимость этих параметров от высоты определяется таблицами из справочника [25]. Вектор $\mathbf{F}_{J_{2,i}}$ – это возмущающее ускорение со стороны второй зональной гармоники J_2 , вычисляемое по формуле

$$\mathbf{F}_{a,i} = -\frac{3}{2} \cdot \mu_E \cdot J_2 \cdot \frac{R_E^2}{R_i^5} \cdot \begin{bmatrix} (1 - 5Z_i^2/R_i^2) X_i \\ (1 - 5Z_i^2/R_i^2) Y_i \\ (3 - 5Z_i^2/R_i^2) Z_i \end{bmatrix}.$$

Система единиц для обезразмеривания уравнений движения была выбрана следующим образом:

$$DU = R_E, \quad VU = \sqrt{\mu_E/R_E}, \quad TU = DU/VU,$$

где DU – это единица расстояния, VU – единица скорости, а TU – единица времени. В такой си-

Таблица 1. Принятые в данной работе численные значения используемых параметров

Параметр	Название	Значение
μ_E	Гравитационный параметр Земли	$398600.44 \text{ км}^3/\text{с}^2$
R_E	Радиус Земли	6371 км
$R_{\text{ref}} - R_E$	Высота круговой орбиты	200 км
J_2	Коэффициент при второй зональной гармонике	$1.0826 \cdot 10^{-3}$
$c_{x,1}$	Коэффициент аэродинамического сопротивления	2.1
A_1/m_1	Отношения площади к массе первого спутника	$0.005-0.020 \text{ м}^2/\text{кг}$
β_2	Баллистический коэффициент второго спутника	$0.01-0.04 \text{ м}^2/\text{кг}$
DU	Единица расстояния	6371 км
VU	Единица скорости	7.910 км/с
TU	Единица времени	13.424 мин

стеме единиц гравитационный параметр Земли равен единице.

В табл. 1 приведены числовые значения всех используемых параметров. Для первого спутника ($i = 1$) задан диапазон, которому принадлежит отношение площади к массе, это значение зависит от ориентации аппарата. Хотя угловое движение аппаратов не моделируется в данной работе, мы считаем отношение площади к массе первого аппарата переменной величиной, которой можно управлять и создавать таким образом различные возмущающие ускорения, действующие на первый аппарат со стороны атмосферы. Для второго аппарата приводится диапазон возможных значений баллистического коэффициента β_2 . Истинное значение этого параметра считается неизвестным. Система управления в режиме реального полета должна будет адаптироваться к этому значению.

Введем орбитальную систему координат $Oxuz$. Начало O поместим на круговую орбиту, радиус которой совпадает с радиусом первого аппарата в некоторый момент времени. Эту орбиту мы будем называть *опорной*. Каждый раз, когда она нам требуется, ее параметры будут пересчитываться на основе параметров орбиты первого спутника и фиксироваться до следующего перерасчета. Положение и скорость движущейся точки O на опорной орбите в инерциальной системе координат мы будем обозначать \mathbf{R}_0 и \mathbf{V}_0 соответственно. Ось Ox направим вдоль радиус-вектора \mathbf{R}_0 , ось Oz направим вдоль вектора $\mathbf{R}_0 \times \mathbf{V}_0$, а ось Oy выберем дополняющей до правой тройки систему $Oxuz$. В орбитальной системе координат положение i -го аппарата обозначим \mathbf{r}_i , а скорость — \mathbf{v}_i .

Фазовый вектор i -го аппарата в инерциальной системе координат обозначим как $\mathbf{X}_i = [\mathbf{R}_i, \mathbf{V}_i]$, а в орбитальной системе координат — как $\mathbf{x}_i = [\mathbf{r}_i, \mathbf{v}_i]$.

Уравнения (1) описывают основную модель движения космических аппаратов в формации, именно в рамках этой модели будет впоследствии моделироваться их движение. Кроме этой модели в работе также используются две другие вспомогательные модели движения — невозмущенная линеаризованная модель относительного движения и возмущенная силой атмосферного сопротивления линеаризованная модель относительного движения. А именно, в рамках невозмущенной линеаризованной модели относительного движения вводится понятие проективной круговой орбиты (PCO), а в рамках возмущенной линеаризованной модели относительного движения строится управление одним из спутников для противодействия дрейфу вдоль орбиты. Опишем эти модели подробнее.

В невозмущенной линеаризованной модели уравнения относительного движения линеаризуются относительно опорной кеплеровой орбиты. В данной работе будем полагать ее круговой, тогда соответствующие уравнения движения спутников называются уравнениями Хилла–Клохесси–Уилтшира и имеют вид:

$$\Delta \ddot{x} - 2n_0 \Delta \dot{y} - 3n_0^2 \Delta x = 0, \quad (2)$$

$$\Delta \ddot{y} + 2n_0 \Delta \dot{x} = 0, \quad (3)$$

$$\Delta \ddot{z} + n_0^2 \Delta z = 0, \quad (4)$$

где $\Delta \mathbf{x}_{12} = \mathbf{x}_2 - \mathbf{x}_1 = [\Delta \mathbf{r}_{12}, \Delta \mathbf{v}_{12}] = [\Delta x, \Delta y, \Delta z, \Delta \dot{x}, \Delta \dot{y}, \Delta \dot{z}]$ — фазовый вектор второго аппарата относительно первого в орбитальной системе координат, а n_0 — среднее движение на опорной орбите. В рамках этой модели считается, что $|\Delta \mathbf{r}_{12}| \ll R_1$. Ограниченные в рамках этой модели траектории отвечают случаю $\Delta \dot{y}(t_0) = -2n_0 \Delta x(t_0)$ (отсутствие дрейфа вдоль орбиты) и принимают следующую форму:

$$\begin{aligned} \Delta x(t) &= \rho_x \sin(n_0(t - t_0) + \alpha_x), \\ \Delta y(t) &= 2\rho_x \cos(n_0(t - t_0) + \alpha_x) + \rho_y, \\ \Delta z(t) &= \rho_z \sin(n_0(t - t_0) + \alpha_z). \end{aligned}$$

Здесь ρ_x , ρ_y , ρ_z , α_x и α_z являются константами, связанными с начальными условиями следующим образом:

$$\rho_x = \frac{1}{n_0} \sqrt{\Delta \dot{x}^2(t_0) + n_0^2 \Delta x^2(t_0)}, \quad (5)$$

$$\rho_y = \Delta y(t_0) - \frac{2\Delta \dot{x}(t_0)}{n_0}, \quad (6)$$

$$\rho_z = \frac{1}{n_0} \sqrt{\Delta \dot{z}^2(t_0) + n_0^2 \Delta z^2(t_0)}, \quad (7)$$

$$\alpha_x = \arctg 2(n_0 \Delta x(t_0), \Delta \dot{x}(t_0)), \quad (8)$$

$$\alpha_z = \arctg 2(n_0 \Delta z(t_0), \Delta \dot{z}(t_0)), \quad (9)$$

где $\arctg 2(\xi, \eta)$ – это функция, которая определяется как угол на плоскости (ξ, η) между положительным направлением оси ξ и лучом из начала координат в точку (ξ, η) . Область значений данной функции примем равной $(-\pi, \pi]$.

Проективной круговой орбитой называют ограниченную в линейном приближении траекторию с $\alpha_x = \alpha_z$ и $\rho_z = 2\rho_x$. Заметим, что в этом случае $2\rho_x = \rho_z = R_{\text{PCO}}$ – величина, которую мы будем называть радиусом проективной круговой орбиты. Для определенности примем $R_{\text{PCO}} = 200$ м.

В рамках возмущенной линеаризованной модели относительного движения в правые части уравнений (2)–(4) добавляются проекции разности $F_{a,2} - F_{a,1}$ сил сопротивления атмосферы, действующих на спутники. В следующем разделе описываются постановка и решение задачи поддержания спутниковой формации в рамках этой модели.

3. ЗАДАЧА ПОДДЕРЖАНИЯ СПУТНИКОВОЙ ФОРМАЦИИ

Задача состоит в синтезе *адаптивного нейрoуправления* орбитальным движением первого аппарата, при котором второй аппарат будет двигаться относительно первого по проективной круговой орбите определенного размера. Предполагается, что данное управление реализуется изменением миделева сечения первого аппарата и с некоторой частотой с помощью серии импульсов. Слово “нейроуправление” означает, что система управления первым аппаратом будет содержать нейронную сеть, а слово “адаптивность” подразумевает, что параметры этой сети будут настраиваться в режиме полета таким образом, чтобы удовлетворить каким-то выбранным изменя-

ющимся условиям. Будем предполагать, что управление вторым аппаратом отсутствует.

Покажем, каким образом задачу синтеза управления для поддержания проективной круговой орбиты можно решить, используя классический подход. Для этого воспользуемся управляемым параметром A_1/m_1 , чтобы с помощью атмосферы по возможности максимально сократить относительный дрейф аппаратов вдоль орбиты. Последующая серия импульсов, совершаемых первым аппаратом, направлена на полное устранение дрейфа и нацеливание на требуемые параметры проективной круговой орбиты.

Плотность атмосферы будем считать постоянной и одинаковой всюду и на всем интервале маневрирования.

Приступим непосредственно к постановке оптимизационной задачи, решение которой и дает требуемое управление для поддержания формации в нелинейной модели с возмущениями. Начнем с поиска значения отношения площади к массе A_1/m_1 у первого аппарата, такого, чтобы максимально сократить дрейф относительно второго аппарата. С одной стороны, разница в больших полуосях первого и второго аппарата в линейном приближении выражается формулой

$$\Delta a_0 = a_2(t_0) - a_1(t_0) = 4\Delta x(t_0) + 2\Delta \dot{y}(t_0)/n_0. \quad (10)$$

С другой стороны, из-за влияния атмосферы, в линейном приближении за период $T_0 = 2\pi/n_0$ опорной орбиты приращения большой полуоси первого и второго аппарата составят соответственно

$$\Delta a_1 = a_1(t_0 + T_0) - a_1(t_0) = -2a_0^2 V_0 F_{a,1} T_0, \quad (11)$$

$$\Delta a_2 = a_2(t_0 + T_0) - a_2(t_0) = -2a_0^2 V_0 F_{a,2} T_0,$$

причем эти формулы выведены в предположении, что возмущающие ускорения $F_{a,1}$ и $F_{a,2}$ постоянны на витке и равны $F_{a,1} = \frac{1}{2} c_{x,1} (A_1/m_1) \rho V_0^2$,

$$F_{a,2} = \frac{1}{2} \beta_2 \rho V_0^2.$$

Условие отсутствия дрейфа через виток вдоль опорной орбиты выглядит следующим образом: $a_1(t_0 + T_0) = a_2(t_0 + T_0)$, что равносильно условию $\Delta a_1 = \Delta a_2 + \Delta a_0$.

Выразим в левой части отношение A_1/m_1 и получим $\frac{A_1}{m_1} = \frac{\Delta a_0 + \Delta a_2}{-a_0^2 V_0^3 T_0 \rho c_{x,1}}$.

Если рассчитанное по этой формуле значение оказывается меньше, чем минимально возможное значение $(A_1/m_1)_{\min}$, то принимаем $(A_1/m_1) = (A_1/m_1)_{\min}$. Если же это значение оказалось больше, чем максимально возможное значение $(A_1/m_1)_{\max}$, то принимаем $(A_1/m_1) = (A_1/m_1)_{\max}$.

Поэтому окончательно формула для отношения площади к массе запишется так:

$$\frac{A_1}{m_1} = \max \left(\left(\frac{A_1}{m_1} \right)_{\min}, \min \left(\frac{\Delta a_0 + \Delta a_2}{-a_0^2 V_0^3 T_0 \rho c_{x,l}}, \left(\frac{A_1}{m_1} \right)_{\max} \right) \right). \quad (12)$$

Обращаем внимание на то, что величина Δa_2 зависит от баллистического коэффициента второго аппарата β_2 , который в режиме реального полета может быть неизвестен. В дальнейшем мы потребуем, чтобы разрабатываемое нейруправление адаптировалось к этому неизвестному значению.

Итак, формула (12) позволяет вычислить отношение площади к массе для первого аппарата, чтобы с помощью атмосферы сократить дрейф второго аппарата. Соответствующее значение дифференциального ускорения торможения в атмосфере вычисляется по формуле $\Delta F = F_{a,2} - F_{a,1}$.

Теперь перейдем к поиску оптимальной серии импульсов для полного устранения дрейфа и попадания на целевую проективную круговую орбиту.

Ограничимся поиском серии четырех импульсов в пределах одного витка. Можно рассмотреть и иное число импульсов, распределенных на иных интервалах времени, но для определенности мы примем эти значения.

Пусть t_1, t_2, t_3 и t_4 – это моменты времени совершения импульсов $\Delta \mathbf{v}_1, \Delta \mathbf{v}_2, \Delta \mathbf{v}_3$ и $\Delta \mathbf{v}_4$ соответственно. По предположению, $t_0 \leq t_1 < t_2 < t_3 < t_4 \leq t_0 + T_0$, где $T_0 = 2\pi/n_0$ – период опорной орбиты, а t_0 – момент начала маневрирования.

В результате применения импульсов скорости $\Delta \mathbf{v}_1, \Delta \mathbf{v}_2, \Delta \mathbf{v}_3$ и $\Delta \mathbf{v}_4$ первым аппаратом в моменты времени t_1, t_2, t_3 и t_4 относительный фазовый вектор второго аппарата с учетом воздействия постоянного возмущающего ускорения со стороны атмосферы в линейном приближении будет равен

$$\begin{aligned} \Delta \mathbf{x}(t_0 + T_0) &= \Phi(t_0 + T_0; t_0) \Delta \mathbf{x}(t_0) - \\ &- \Phi(t_0 + T_0 - t_1; t_0) \delta \mathbf{x}_1 - \Phi(t_0 + T_0 - t_2; t_0) \delta \mathbf{x}_2 - \\ &- \Phi(t_0 + T_0 - t_3; t_0) \delta \mathbf{x}_3 - \Phi(t_0 + T_0 - t_4; t_0) \delta \mathbf{x}_4 + \\ &+ \Delta F/n_0 \cdot [4\pi/n_0, -6\pi^2/n_0, 0, 0, -6\pi, 0], \end{aligned}$$

где $\delta \mathbf{x}_i = [0, 0, 0, \Delta \mathbf{v}_i]$, а матрица $\Phi(\tau_1; \tau_0)$ – это матрица перехода линеаризованной системы уравнений (2)–(4), она имеет следующий вид:

$$\begin{aligned} &\Phi(\tau_1; \tau_0) = \\ &= \begin{bmatrix} 4 - 3c & 0 & 0 & s/n_0 & 2(1-c)/n_0 & 0 \\ 6(s - \Delta u) & 1 & 0 & 2(c-1)/n_0 & (4s - 3\Delta u)/n_0 & 0 \\ 0 & 0 & c & 0 & 0 & s/n \\ 3n_0 s & 0 & 0 & c & 2s & 0 \\ 6n_0(c-1) & 0 & 0 & -2s & 4c-3 & 0 \\ 0 & 0 & -n_0 s & 0 & 0 & c \end{bmatrix}, \end{aligned}$$

где $\Delta u = n_0(\tau_1 - \tau_0)$, $s = \sin \Delta u$, $c = \cos \Delta u$.

Теперь, зная $\Delta \mathbf{x}(t_0 + T_0)$, по формулам (5) можно вычислить параметры относительной орбиты $\rho_x, \rho_y, \rho_z, \alpha_x, \alpha_z$ и сравнить их с требуемыми параметрами. Это дает четыре уравнения

$$\rho_x = R_{\text{PCO}}/2, \quad \rho_y = 0, \quad \rho_z = R_{\text{PCO}}, \quad \alpha_x = \alpha_z, \quad (13)$$

к которым следует также добавить условие отсутствия дрейфа

$$\Delta \dot{y}(t_0 + T_0) + 2n_0 \Delta x(t_0 + T_0) = 0. \quad (14)$$

В результате получаются 5 уравнений, играющих роль ограничений-равенств в формулируемой оптимизационной процедуре. Ограничениями-неравенствами служат неравенства

$$t_0 < t_1, \quad t_1 < t_2, \quad t_2 < t_3, \quad t_3 < t_4, \quad t_4 < t_0 + T_0. \quad (15)$$

В качестве оптимизируемого функционала мы выбрали квадратичный функционал $J = \sum_{i=1}^4 |\Delta \mathbf{v}_i|^2$.

Итак, мы привели основанное на классических подходах решение задачи о поддержании проективной круговой орбиты в линейной модели с возмущениями от силы сопротивления атмосферы. Именно в этих предположениях будет происходить обучение управляющих нейронных сетей, о которых подробно пойдет речь в следующем разделе. Моделирование работы этих нейронных сетей и управления будет производиться в основной нелинейной модели движения (1).

4. УПРАВЛЯЮЩИЕ НЕЙРОННЫЕ СЕТИ

Резюмируем в виде алгоритма предложенный в предыдущем разделе метод расчета всех необходимых для управления формацией параметров.

1. В момент времени t_0 даны расстояние до первого аппарата R_1 , его скорость V_1 , разница фазовых векторов аппаратов в орбитальной системе координат $\Delta \mathbf{x}_{12} = \mathbf{x}_2 - \mathbf{x}_1$, плотность атмосферы в точке, где расположен первый аппарат ρ , и баллистический коэффициент второго аппарата β_2 .

2. На основании этих данных вычисляются большая полуось опорной орбиты a_0 , скорость движения по опорной орбите V_0 , среднее движение для опорной орбиты n_0 , ее период T_0 , невязка по большим полуосям аппаратов Δa_0 (10), изменение большой полуоси второго аппарата за период опорной орбиты Δa_2 (11), отношение площади к массе первого аппарата A_1/m_1 (12) и дифференциальное ускорение торможения ΔF . Величины $\Delta a_2, A_1/m_1$ и ΔF зависят от баллистического коэффициента β_2 второго аппарата.

3. Ставится задача оптимизации целевой функции J_1 или J_2 при ограничениях-равенствах (13), (14) и ограничениях-неравенствах (15). В резуль-

тате решения оптимизационной задачи определяются моменты времени t_1, t_2, t_3, t_4 совершения импульсов скорости и сами импульсы $\Delta v_1, \Delta v_2, \Delta v_3, \Delta v_4$.

Оптимизационная задача решается с помощью метода последовательного квадратичного программирования, включающего метод активных множеств и линейный поиск. Для простоты моменты времени было решено зафиксировать и считать $n_0 t_1 = \pi/6, n_0 t_2 = 2\pi/3, n_0 t_3 = 4\pi/3, n_0 t_4 = 11\pi/6$, численные эксперименты показали, что выбор конкретных моментов времени слабо влияет на величину оптимизируемого функционала. Выбор других значений также возможен. Для фиксированных моментов времени необходимость в ограничениях-неравенствах отпадает.

Для управления движением первого аппарата предлагается использовать две управляющие сети. Одна сеть аппроксимирует отображение $[R_1, V_1, \Delta x, \rho, \beta_2] \rightarrow [\Delta v_1, \Delta v_2, \Delta v_3, \Delta v_4]$ и будет обозначаться как функция $N_{\Delta v} = N_{\Delta v}(s)$, $s = [R_1, V_1, \Delta x, \rho, \beta_2]$. Эта функция принимает на вход 10 параметров и выдает 12 параметров. Вторая управляющая сеть аппроксимирует отображение $[R_1, V_1, \Delta x, \rho, \beta_2] \rightarrow A_1/m_1$ и будет обозначаться как функция $N_{A/m} = N_{A/m}(s)$. Эта функция принимает на вход 10 параметров и выдает 1 параметр. Аппроксиматор отношения A_1/m_1 выделяется в отдельную нейронную сеть для того, чтобы, с одной стороны, упростить аппроксимируемые отображения и тем самым повысить скорость их обучения, а также для того, чтобы выбрать ограниченную активационную функцию на выходном слое этой нейронной сети, так как отношение A_1/m_1 может принимать значения в определенном интервале.

Выборка для обучения нейронных сетей $N_{\Delta v}$ и $N_{A/m}$ формировалась следующим образом. Сначала были случайным образом сформированы значения для входных переменных:

1. $R_1 \in U(R_{\text{ref}} - \Delta R, R_{\text{ref}} + \Delta R), \Delta R = 10 \text{ км},$
2. $V_1 \in U(V_{\text{ref}} - \Delta V, V_{\text{ref}} + \Delta V), \Delta V = 10 \text{ м/с},$
3. $\Delta r_{12} \in U([- \Delta r, \Delta r]^3), \Delta r = 200 \text{ м},$
4. $\Delta v_{12} \in U([- \Delta v, \Delta v]^3), \Delta v = 0.2 \text{ м/с},$
5. $\rho \in U(10^{-12}, 10^{-11}) \text{ кг/м}^3,$
6. $\beta_2 \in U(0.01, 0.04) \text{ м}^2/\text{кг}.$

Здесь $U(a, b)$ означает равномерное распределение на интервале $[a, b]$, $U([a, b]^n)$ означает равномерное распределение в гиперкубе $[a, b]^n$. Для каждой реализации случайных величин было вы-

числено значение A_1/m_1 и найдены оптимальные значения $\Delta v_1, \Delta v_2, \Delta v_3, \Delta v_4$.

Начальные приближения для импульсов скорости выбирались нулевыми векторами. В некоторых случаях оптимизационная процедура не сходилась. В обучающую выборку попадали только те выходы оптимизационной процедуры, которые сигнализируют о сходимости процесса, т.е. когда градиент функции Лагранжа по модулю не превышал 10^{-6} . В результате была сформирована выборка из 200 тысяч пар вход/выход.

Для обучения нейронных сетей использовалась библиотека Deep Learning Toolbox (Version 13.0) в среде Matlab (R2019b). Были проанализированы как неглубокие, так и глубокие нейронные сети прямого распространения с различным числом нейронов на скрытых слоях. Чтобы отличать нейронные сети с различным числом нейронов, мы будем называть нейронной сетью размера $m \times n$ нейронную сеть, у которой n скрытых слоев и в каждом скрытом слое m нейронов. Графы, которые образуют нейронные сети, являются полностью связными. На всех скрытых слоях активационной функцией был гиперболический тангенс $\text{th}(x)$. Другие широко применяемые активационные функции $\max(x, 0)$ и радиально-базисная функция во всех случаях дали худшую производительность и качество нейронной сети, чем функция $\text{th}(x)$. На выходном слое нейронной сети $N_{\Delta v}$ активационная функция является тождественной функцией $\phi(x) = x$, чтобы не вводить искусственным образом ограничения на выход. На выходном слое нейронной сети $N_{A/m}$ взята активационная функция $\text{th}(x)$, так как эта величина имеет физические ограничения.

Обучение нейронных сетей выполнялось с помощью метода масштабируемых сопряженных градиентов (SCG). В качестве целевого функционала выбрано среднеквадратическое отклонение (MSE) выхода нейронных сетей от целевых значений. Перед обучением выборка была предварительно масштабирована таким образом, чтобы значения входных и целевых параметров лежали в интервале $[-1, 1]$. Выборка делилась на три части – обучающую, валидационную и тестовую – в соотношении 70/15/15 (обучающая выборка содержит 140000 элементов, валидационная и тестовая выборки – по 30000 элементов). Обучение сетей прекращалось только в случае, когда улучшение функционала на валидационной части отсутствовало 50 эпох подряд.

Кроме MSE для оценки качества нейронных сетей $N_{\Delta v}$ различного размера удобно использовать следующий показатель:

$$Q_{N_{\Delta v}} = Q_1 + Q_2 + Q_3 + Q_4, Q_j = \frac{100\%}{|I_t|} \sum_{i \in I_t} \mathbf{I} \left(\frac{|(\Delta v_j^p)_i - (\Delta v_j^t)_i|}{|\Delta v_j^t|} > 0.1 \right),$$

где индекс $j = 1, 2, 3, 4$ нумерует импульс скорости, I_t – множество индексов, отвечающее тестовой части выборки в общей выборке, $|I_t|$ – обозначает число элементов в множестве I_t , i – номер примера из тестовой выборки, \mathbf{I} – означает индикатор события, p – означает величину на выходе нейронной сети, t – означает целевое значение из тестовой выборки, а под нормой понимается обычная евклидова норма. По сути, величины Q_j представляют собой доли случаев, когда относительная погрешность определения импульса превышает 10%. Величина $Q_{N_{\Delta v}}$ суммирует эти доли. В идеальном случае, если $Q_j = 0\%$, это будет означать, что относительная погрешность определения Δv_j в каждом тестовом примере не превышает 10%. В худшем случае $Q_j = 100\%$, это значит, что относительная погрешность определения Δv_j во всех тестовых примерах превысила 10%.

В табл. 2 приводятся результаты обучения нейронных сетей для различных архитектур. Также в таблице указаны значения MSE_{tr} , MSE_v , MSE_{test} среднеквадратичных ошибок, подсчитанных на обучающей, валидационной и тестовой частях выборки соответственно. Горизонтальная черта отделяет результаты для сетей с различным числом скрытых слоев.

Лучшая нейронная сеть отбиралась по следующему правилу. Во-первых, из рассмотрения были удалены все “дефектные” и переобученные нейронные сети, т.е. все сети, у которых $MSE_{test}/MSE_{tr} < 0.1$ и $MSE_{test}/MSE_{tr} > 10$. Во-вторых, были удалены нейронные сети, для которых существуют другие нейронные сети одновременно с меньшими значениями ошибок MSE_{tr} , MSE_v , MSE_{test} на обучающей, валидационной и тестовой частях выборки соответственно. В результате осталось пять нейронных сетей размера 100×2 , 60×3 , 80×3 , 90×3 и 100×3 (в табл. 2 соответствующие этим сетям строки выделены цветами). Далее мы объединили валидационную и тестовую части выборки (назовем ее тоже тестовой частью, а соответствующее значение MSE обозначим за MSE_{test}) и для каждой нейронной сети получили значения MSE_{tr} и MSE_{test} . Мы исключили нейронную сеть размера 100×3 , так как по сравнению с другими сетями MSE_{tr} у нее был незаметно лучше, а MSE_{test} сильно хуже. В результате мы оставили только сеть размера 90×3 , так как для нее оба значения MSE_{tr} и MSE_{test} оказались меньше, чем у остальных сетей.

Итак, в результате анализа среднеквадратичных ошибок мы остановились на сети размера 90×3 . Именно эта сеть и будет участвовать по всех дальнейших расчетах. Судя по табл. 2, в 2.82% тестовых случаев относительная ошибка исполнения первого импульса превысила 10%. В 12.86% случаев относительная ошибка исполнения второго импульса превысила 10%. В 15.81% случаев относительная ошибка исполнения третьего импульса превысила 10%. И, наконец, в 8.57% случаев относительная ошибка исполнения четвертого импульса превысила 10%.

Что касается сети $N_{A/m}$, ее качество, помимо MSE_{tr} , MSE_v , MSE_{test} , оценивалось аналогичным образом:

$$Q_{N_{A/m}} = \frac{100\%}{|I_t|} \sum_{i \in I_t} \mathbf{I} \left(\frac{|(A_1/m_1)_i^p - (A_1/m_1)_i^t|}{|(A_1/m_1)_i^t|} > 10^{-5} \right),$$

где индекс i пробегает все тестовые элементы выборки, индекс p означает выход нейронной сети, t означает целевое значение. Результаты обучения нейронных сетей $N_{A/m}$ различных размеров приведены в табл. 3.

Лучшая нейронная сеть отбиралась тем же способом, что и сеть $N_{\Delta v}$: сначала были удалены дефектные и переобученные сети, потом все сети, для которых были найдены сети лучше по всем показателям MSE_{tr} , MSE_v , MSE_{test} . Объединять тестовую и валидационную выборку не пришлось, в результате этих двух шагов была обнаружена лучшая сеть – она имеет размер 80×3 . Как следует из табл. 3, для этой сети относительная ошибка определения A_1/m_1 на тестовой части выборки превысила 0.001% только в 3.54% случаев.

5. АДАПТАЦИЯ НЕЙРОННЫХ СЕТЕЙ

Перейдем к процессу адаптации управления к различным значениям баллистического коэффициента β_2 второго аппарата.

Движение космических аппаратов моделируется согласно уравнениям движения (1). Процесс моделирования состоит из двух циклов: внешнего (эпизоды) и внутреннего (шаги). Каждый эпизод начинался со случайно сгенерированного состояния $s = [R_1, V_1, \Delta x, \rho, \beta_2]$, где значение баллистического коэффициента β_2 является оценкой истинного и неизвестного для первого аппарата баллистического коэффициента β_2^{true} второго аппарата. Заметим, что плотность атмосферы ρ мы берем здесь случайной величиной, равномерно распределенной на интервале $[10^{-12}, 10^{-11}]$ кг/м³. Это позволит нам, с одной стороны, не привязываться к конкретной модели атмосферы, а с другой сторо-

Таблица 2. Качество нейронных сетей $N_{\Delta v}$ различных размеров. Цветами выделены отобранные недефектные и непереобученные сети. По результатам анализа выбрана сеть размера 90×3 , соответствующая строка выделена оранжевым цветом

$m \times n$	$MSE_{tr}, 10^{-5}$	$MSE_v, 10^{-5}$	$MSE_{test}, 10^{-5}$	$Q_1, \%$	$Q_2, \%$	$Q_3, \%$	$Q_4, \%$	$Q, \%$
10 × 1	7.19	7.60	7.15	10.80	53.75	58.13	41.99	164.67
20 × 1	5.87	6.44	6.25	9.49	45.20	44.48	28.24	127.40
30 × 1	4.60	4.62	4.19	4.83	23.80	26.24	14.14	69.01
40 × 1	6.00	6.12	5.91	9.06	43.41	42.40	27.63	122.49
50 × 1	4.22	4.13	4.25	4.12	17.66	21.64	10.95	54.37
60 × 1	4.22	4.35	4.05	4.12	19.40	22.81	10.82	57.15
70 × 1	3.89e	4.21	4.28	3.72	17.32	20.73	10.74	52.50
80 × 1	4.27e	4.01	4.22	4.58	19.82	24.07	11.60	60.06
90 × 1	4.14e	4.22	4.15	4.13	16.75	21.51	10.48	52.87
100 × 1	4.61	4.11	4.26	4.86	22.37	26.23	13.25	66.71
10 × 2	5.86	5.67	5.53	7.65	45.29	51.98	34.17	139.09
20 × 2	4.64	5.13	4.91	5.26	27.89	44.98	25.80	103.93
30 × 2	3.83	4.07	3.32	3.56	16.09	20.19	10.22	50.07
40 × 2	3.46	3.95	3.79	3.15	14.29	18.50	8.87	44.81
50 × 2	3.40	3.30	3.57	3.11	13.50	16.47	8.58	41.66
60 × 2	3.65	3.74	3.55	3.46	16.09	21.92	11.94	53.40
70 × 2	3.44	3.50	3.34	3.39	16.86	22.04	10.48	52.77
80 × 2	3.17	3.31	3.18	2.91	11.69	14.08	7.95	36.64
90 × 2	3.17	3.00	3.24	2.54	10.55	13.15	7.53	33.77
100 × 2	2.95	3.30	2.93	1.99	9.11	11.45	7.10	29.65
10 × 3	4.94	5.11	4.43	6.83	30.40	43.82	25.86	106.91
20 × 3	3.83	3.78	4.22	3.28	16.06	35.99	22.37	77.70
30 × 3	3.50	3.84	3.44	3.45	16.75	19.34	10.27	49.81
40 × 3	3.49	3.28	3.23	2.88	13.73	16.72	8.04	41.37
50 × 3	3.02	3.08	3.12	3.00	13.38	16.90	8.93	42.21
60 × 3	2.97	3.17	3.03	3.12	14.04	18.02	9.21	44.39
70 × 3	3.26	3.42	3.54	3.38	14.53	19.87	9.10	46.87
80 × 3	3.19	3.24	2.94	3.08	14.61	18.17	9.49	45.34
90 × 3	2.88	2.86	3.10	2.79	12.86	15.81	8.57	40.02
100 × 3	2.82	3.17	3.43	2.82	13.56	16.49	8.72	41.59

ны, – протестировать нейроуправление в более сложных условиях.

После инициализации состояния системы s шла последовательность шагов, на каждом из которых в течение одного витка выполнялось маневрирование первого аппарата: управляющие нейронные сети выдавали требуемые A_1/m_1 и импульсы скорости. Очередное маневрирование выполнялось сразу после завершения предыдущего маневрирования, хотя можно включить и пассивное движение аппаратов после маневрирования. В начале каждого шага параметры опорной орбиты пересчитывались: большая полуось, наклоне-

ние, долгота восходящего узла и истинная долгота брались равными соответствующим значениям для первого аппарата в конце интервала маневрирования, а эксцентриситет опорной орбиты полагался равным нулю.

Истинное значение баллистического коэффициента β_2^{true} второго аппарата неизвестно первому аппарату. На вход β_2 для управляющих нейронных сетей подавалось фиксированное (тестовое) значение. Для того чтобы понять, насколько хорошо подходит выбранное тестовое значение β_2 , мы накапливали статистику – евклидову норму векто-

Таблица 3. Качество нейронных сетей $N_{A/m}$ различных размеров. Оранжевым цветом выделена строка с сетью 80×3 , отобранной по результатам анализа качества

$m \times n$	$MSE_{tr}, 10^{-6}$	$MSE_v, 10^{-6}$	$MSE_{test}, 10^{-6}$	$Q_{N_{A/m}}, \%$
10×1	206	213	204	11.51
20×1	181	193	187	9.05
30×1	199	216	190	11.83
40×1	192	209	219	11.81
50×1	192	214	203	11.56
60×1	191	207	192	10.61
70×1	197	220	220	12.08
80×1	192	215	218	12.50
90×1	209	221	224	12.02
100×1	186	227	234	12.02
10×2	34.4	36.7	35.6	5.58
20×2	11.4	12.6	14.1	4.17
30×2	3.30	3.80	4.38	3.43
40×2	16.7	18.7	18.6	4.36
50×2	2.66	4.03	5.38	3.53
60×2	23.9	31.3	32.7	4.91
70×2	6.86	9.46	10.2	3.54
80×2	5.27	20.5	19.2	3.85
90×2	21.0	36.6	32.1	4.41
100×2	17.6	29.8	32.6	4.25
10×3	187	191	186	12.88
20×3	2.91	3.69	4.14	3.80
30×3	5.36	8.26	7.48	3.77
40×3	4.73	11.5	7.39	3.95
50×3	1.31	4.10	4.71	3.38
60×3	4.38	10.9	10.9	4.09
70×3	0.163	1.64	1.98	3.03
80×3	1.98	8.36	13.5	3.54
90×3	0.0238	0.992	2.58	2.95
100×3	1.43	12.6	9.98	3.50

ра $\Delta = [\rho_x - R_{PCO}/2, \rho_y, \rho_z - R_{PCO}, \sin \alpha_x - \sin \alpha_z, \cos \alpha_x - \cos \alpha_z]$, где $\rho_x, \rho_y, \rho_z, \alpha_x, \alpha_z$ – параметры относительной орбиты, на которой оказался второй аппарат в результате маневрирования первого аппарата. Эти числа рассчитывались по формулам (5).

Итак, пусть в результате моделирования эпизодов и шагов была собрана выборка из $|\Delta_i|$, $i = 1, \dots, K$, объема K . Мерой качества управления с выбранным значением β_2 мы считаем выборочное среднее $\Delta(\beta_2) = \frac{1}{K} \sum_{i=1}^K |\Delta_i(\beta_2)|$.

Задача адаптации управления к истинному значению β_2 ставится как одномерная оптимизационная задача $\Delta \rightarrow \min_{\beta_2}$. Предполагается, что аппарат будет решать эту оптимизационную задачу в режиме полета, поэтому схема оптимизации должна быть по возможности наиболее простой с вычислительной точки зрения.

Численные эксперименты показали, что эту задачу можно решать неградиентным методом оптимизации Хука–Дживса (pattern search, [26]). Ниже приводится алгоритм оптимизационной процедуры.

1. Выбрать какое-либо значение β_2 из области значений, на которой обучались нейронные сети. В нашем случае это был интервал $[0.01, 0.04] \text{ м}^2/\text{кг}$.

2. Провести моделирование полета аппаратов с фиксированным значением β_2 на входе в управляющие нейронные сети. Подавать на вход управляющим нейронным сетям случайное значение плотности атмосферы, распределенное равномерно на интервале $[10^{-12}, 10^{-11}] \text{ кг/м}^3$. Сформировать выборку из $|\Delta_i|$, $i = 1, \dots, K$, невязок по параметрам относительной орбиты и вычислить выборочное среднее $\Delta(\beta_2) = \frac{1}{K} \sum_{i=1}^K |\Delta_i(\beta_2)|$.

3. Выполнить пункт 2 для соседних значений $\beta_2 - \delta\beta_2$ и $\beta_2 + \delta\beta_2$, где δ_2 можно выбрать, например, равной 0.01. Вычислить таким образом значения $\Delta(\beta_2 - \delta\beta_2)$ и $\Delta(\beta_2 + \delta\beta_2)$.

4. В качестве нового значения β_2 выбрать то значение из $\beta_2 - \delta\beta_2$, β_2 и $\beta_2 + \delta\beta_2$, для которого функционал Δ оказался минимальным. Если таковым оказалось значение $\Delta(\beta_2)$, то шаг $\delta\beta_2$ уменьшается, например, в 2 раза и пункты 3 и 4 выполняются заново. Иначе перейти к пункту 2.

Расчеты прекращаются, если $\delta\beta_2 < \epsilon$, где ϵ можно взять, например, равным $10^{-3} \text{ м}^2/\text{кг}$. Заметим, что в случаях, когда $\delta\beta_2$ не меняется, а β_2 все время только увеличивается или только уменьшается, то получаются значения функционала Δ на равномерной сетке. Память значений на этой сетке функционала позволяет не проводить моделирование в уже рассмотренных узлах сетки.

На рис. 1 приведен график функции $\Delta(\beta_2)$ на равномерной сетке из 100 значений. Число используемых для расчета эпизодов и шагов равнялось 50. Негладкость функции объясняется тем, что в процессе моделирования используются случайные величины.

В табл. 4 приводятся итерации метода оптимизации для случая, когда $\beta_2^{\text{true}} = 0.03 \text{ м}^2/\text{кг}$, начальное значение $\beta_2 = 0.015 \text{ м}^2/\text{кг}$, начальный шаг $\delta\beta_2 = 0.01 \text{ м}^2/\text{кг}$, критерием остановки служило

значение $\epsilon = 0.001 \text{ м}^2/\text{кг}$, число эпизодов равно 30, число шагов в каждом эпизоде равно 20. В случаях, когда необходимо было обновить шаг $\delta\beta_2$, он уменьшался в 2 раза. Всего было выполнено 8 итераций. Табл. 4 показывает, как быстро происходит сходимость процедуры: после первой итерации получается решение, которое отличается от точного ($0.03 \text{ м}^2/\text{кг}$) на $0.005 \text{ м}^2/\text{кг}$. После 4 итераций решение отличается от точного на $0.0025 \text{ м}^2/\text{кг}$.

На рис. 2 показано распределение нормы вектора невязки по параметрам относительной орбиты, когда оптимизация проводилась для случая $\beta_2 = 0.015 \text{ м}^2/\text{кг}$ (до адаптации) и $\beta_2 = 0.0269 \text{ м}^2/\text{кг}$ (после адаптации). Гистограммы же построены для более объемной выборки – когда число эпизодов и число шагов в каждом из них равно 50. Видно, что доля случаев, когда невязка велика (скажем, 0.25), уменьшилась, а доля случаев, когда невязка мала, – увеличилась.

Число эпизодов и число шагов являются свободными параметрами метода. Эти числа должны быть достаточно велики, чтобы получить репрезентативную выборку невязок по параметрам относительной орбиты и устойчивое значение среднего Δ . Впрочем, выбор меньших их значений тоже может привести к решению, близкому к истинному значению β_2^{true} . Например, в табл. 5 можно найти примеры итераций для случая, когда число эпизодов и число шагов в каждом из них равно 10. Это значит, что для сбора статистики и оценки Δ требуется порядка 100 витков вдоль опорной орбиты, т.е. порядка 6.4 дней. Учитывая, что число итераций метода равно 8 и на каждой итерации приходится 1–2 раза собирать статистику, то на адаптацию в целом будет уходить от 50 до 100 дней. Рис. 3 показывает, как меняется распределение невязок по параметрам относительной орбиты до и после адаптации в этом случае. Как и раньше, гистограммы построены для

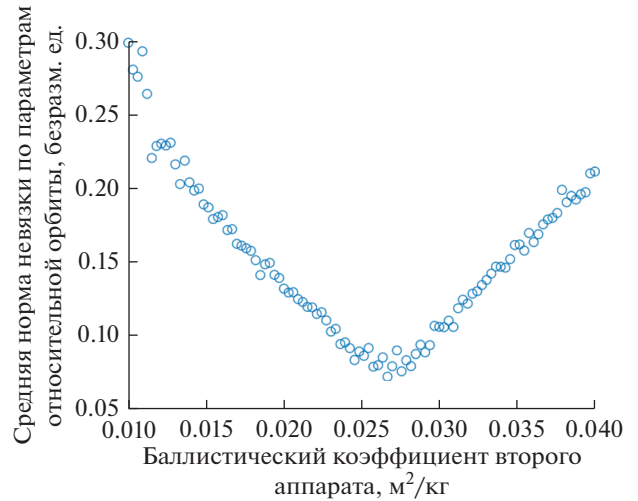


Рис. 1

более объемной выборки – когда число эпизодов и число шагов в каждом из них равно 50.

Гистограммы, показанные на рис. 2 и 3, очень похожи, это объясняется тем, что найденные оптимальные значение β_2 в обоих случаях оказались одинаковыми. Разница имеет место лишь потому, что во время оптимизации при моделировании были разные последовательности реализаций случайных величин, так как число эпизодов и число шагов отличались.

Отметим, наконец, что адаптация стала возможной даже для случая, когда рассматривался только один эпизод и 5 шагов. Результаты расчетов приводятся в табл. 6 и на рис. 4. Таким образом, для накопления статистики было достаточно сделать 5 витков по орбите. Это значит, что на весь процесс адаптации будет уходить порядка 2.5–5 дней. Кроме того, заметим, что малое число витков для накопления статистики может быть

Таблица 4. Итерации во время адаптации значения β_2 в случае, когда $\beta_2^{\text{true}} = 0.03 \text{ м}^2/\text{кг}$, начальное значение $\beta_2 = 0.015 \text{ м}^2/\text{кг}$, начальный шаг $\delta\beta_2 = 0.01 \text{ м}^2/\text{кг}$. Критерием остановки служило значение $\epsilon = 0.001 \text{ м}^2/\text{кг}$, число эпизодов равно 30, число шагов в каждом эпизоде равно 20

Номер итерации	$\beta_2, \text{ м}^2/\text{кг}$	$\delta\beta_2, \text{ м}^2/\text{кг}$	$\Delta(\beta_2 - \delta\beta_2)$	$\Delta(\beta_2)$	$\Delta(\beta_2 + \delta\beta_2)$
1	0.0150	0.0100	0.38645	0.19621	0.09407
2	0.0250	0.0100	0.19631	0.09407	0.16965
3	0.0250	0.0050	0.15086	0.09407	0.10525
4	0.0250	0.0025	0.12432	0.09407	0.08629
5	0.0275	0.0025	0.10068	0.08629	0.10016
6	0.0275	0.0013	0.08722	0.08629	0.10426
7	0.0275	0.0006	0.08597	0.08629	0.09406
8	0.0269	0.0006	0.08810	0.08597	0.09344

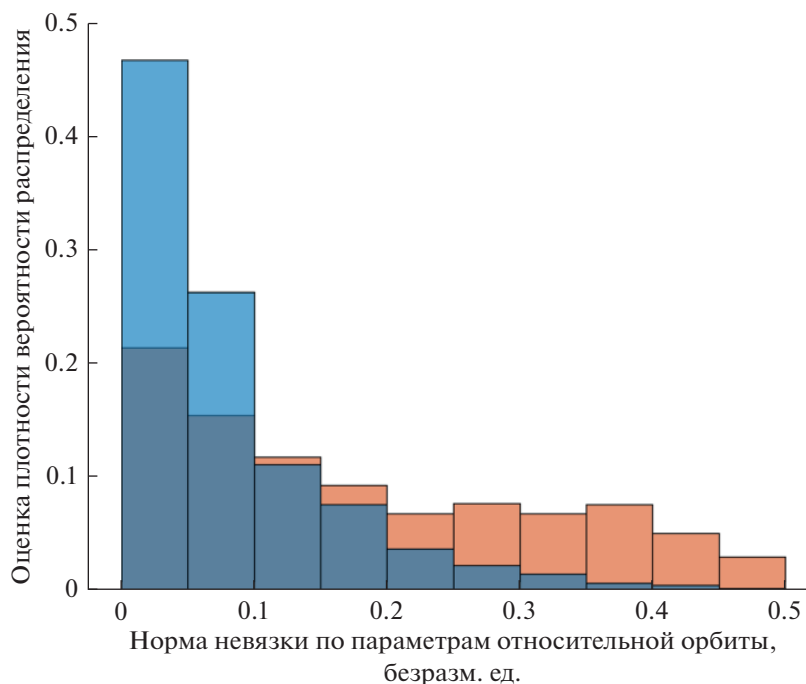


Рис. 2. Во время оптимизации число эпизодов равно 30, число шагов в каждом эпизоде равно 20. Оценка распределения до оптимизации (красное) и после оптимизации (синяя) выполняется на 50 эпизодах и 50 шагах в каждом эпизоде.

особенно важно на раннем этапе адаптации, чтобы как можно раньше попасть в область значений β_2 , близких к истинному значению.

Качество полета можно улучшить, адаптируясь также к истинному неизвестному значению плотности атмосферы ρ^{true} . Будем для простоты считать, что ρ^{true} не зависит от времени, а на вход управляющим нейронным сетям подается тестовое фиксированное значение плотности атмосферы ρ . В таком случае функционал $\Delta = \Delta(\beta_2, \rho)$ становится функцией уже двух переменных. За-

дача адаптации управления к истинным значениям β_2 и ρ ставится как двумерная задача оптимизации $\Delta \rightarrow \min_{\beta_2, \rho}$. Эту задачу можно решать тем же неградиентным методом оптимизации Хука–Дживса, но в этом случае решение ищется уже не на отрезке, а на некотором прямоугольнике в двумерном пространстве. Алгоритм, приведенный выше для одномерной оптимизации, естественным образом обобщается на двумерный (и на любой многомерный) случай, для этого в окрестности каждого вектора (β_2, ρ) на каждой итерации рассматриваются векторы $(\beta_2 - \delta\beta_2, \rho)$, $(\beta_2 + \delta\beta_2, \rho)$,

Таблица 5. Итерации во время адаптации значения β_2 в случае, когда $\beta_2^{\text{true}} = 0.03 \text{ м}^2/\text{кг}$, начальное значение $\beta_2 = 0.015 \text{ м}^2/\text{кг}$, начальный шаг $\delta\beta_2 = 0.01 \text{ м}^2/\text{кг}$. Критерием остановки служило значение $\epsilon = 0.001 \text{ м}^2/\text{кг}$, число эпизодов равно 10, число шагов в каждом эпизоде равно 10

Номер итерации	$\beta_2, \text{ м}^2/\text{кг}$	$\delta\beta_2, \text{ м}^2/\text{кг}$	$\Delta(\beta_2 - \delta\beta_2)$	$\Delta(\beta_2)$	$\Delta(\beta_2 + \delta\beta_2)$
1	0.0150	0.0100	0.36112	0.21124	0.08654
2	0.0250	0.0100	0.19902	0.08654	0.16229
3	0.0250	0.0050	0.17726	0.08654	0.13510
4	0.0250	0.0025	0.10291	0.08654	0.08343
5	0.0275	0.0025	0.08756	0.08343	0.09498
6	0.0275	0.0013	0.08434	0.08343	0.09998
7	0.0275	0.0006	0.07863	0.08343	0.09430
8	0.0269	0.0006	0.08785	0.07863	0.09508

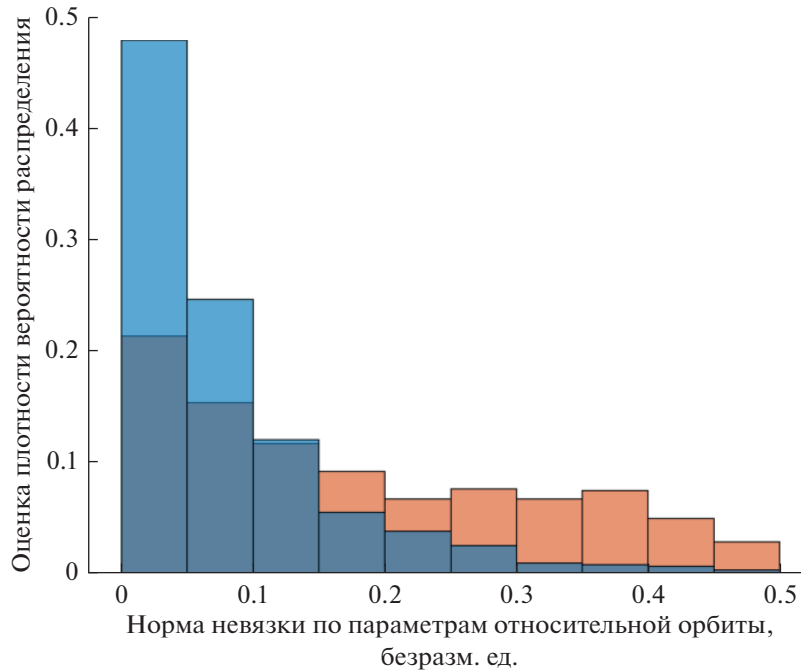


Рис. 3. Во время оптимизации число эпизодов равно 10, число шагов в каждом эпизоде равно 10. Оценка распределения до и после оптимизации как на рис. 2.

$(\beta_2, \rho - \delta\rho)$, $(\beta_2, \rho + \delta\rho)$, где, как и раньше, $\delta\beta_2$ – шаг по значениям β_2 , а $\delta\rho$ – шаг по значениям ρ .

На рис. 5 изображены значения $\Delta(\beta_2, \rho_2)$ нормы невязки по параметрам относительной орбиты для различных значений параметров β_2 и ρ_2 . Значения $\Delta(\beta_2, \rho_2)$ рассчитывались для случая, когда $\beta_2^{\text{true}} = 0.03 \text{ м}^2/\text{кг}$, $\rho^{\text{true}} = 8 \cdot 10^{-12} \text{ кг}/\text{м}^3$, число эпизодов и число шагов в каждом эпизоде равнялось 50. Хорошо видно, что минимальные значения этой функции сосредоточены в окрестности истинных значений параметров β_2 и ρ_2 .

На рис. 6 показано распределение нормы невязки по параметрам относительной орбиты до и после

адаптации для случая, когда число эпизодов и шагов равно 10. Моделирование проводилось для значений $\beta_2^{\text{true}} = 0.03 \text{ м}^2/\text{кг}$, $\rho^{\text{true}} = 8 \cdot 10^{-12} \text{ кг}/\text{м}^3$, начальные приближения для оптимизационной процедуры выбраны равными $\beta_2 = 0.015 \text{ м}^2/\text{кг}$, $\rho = 5 \cdot 10^{-12} \text{ кг}/\text{м}^3$. Для верификации результатов гистограммы построены в результате моделирования 50 эпизодов и 50 шагов в каждом эпизоде. Видно, что адаптация существенно повышает точность управления и, более того, она становится даже выше, чем в случаях, когда плотность атмосферы не адаптируется (ср. с рис. 3).

Следует отметить, что адаптация одновременно двух параметров требует в 2–4 раза больше

Таблица 6. Итерации во время адаптации значения β_2 в случае, когда $\beta_2^{\text{true}} = 0.03 \text{ м}^2/\text{кг}$, начальное значение $\beta_2 = 0.015 \text{ м}^2/\text{кг}$, начальный шаг $\delta\beta_2 = 0.01 \text{ м}^2/\text{кг}$. Критерием остановки служило значение $\epsilon = 0.001 \text{ м}^2/\text{кг}$, число эпизодов равно 1, число шагов в каждом эпизоде равно 5

Номер итерации	$\beta_2, \text{ м}^2/\text{кг}$	$\delta\beta_2, \text{ м}^2/\text{кг}$	$\Delta(\beta_2 - \delta\beta_2)$	$\Delta(\beta_2)$	$\Delta(\beta_2 + \delta\beta_2)$
1	0.0150	0.0100	0.32831	0.22958	0.12593
2	0.0250	0.0100	0.18645	0.12593	0.35300
3	0.0250	0.0050	0.13368	0.12593	0.32613
4	0.0250	0.0025	0.05039	0.12593	0.02982
5	0.0275	0.0025	0.08887	0.02982	0.07011
6	0.0275	0.0013	0.17544	0.02982	0.14247
7	0.0275	0.0006	0.09733	0.02982	0.04134

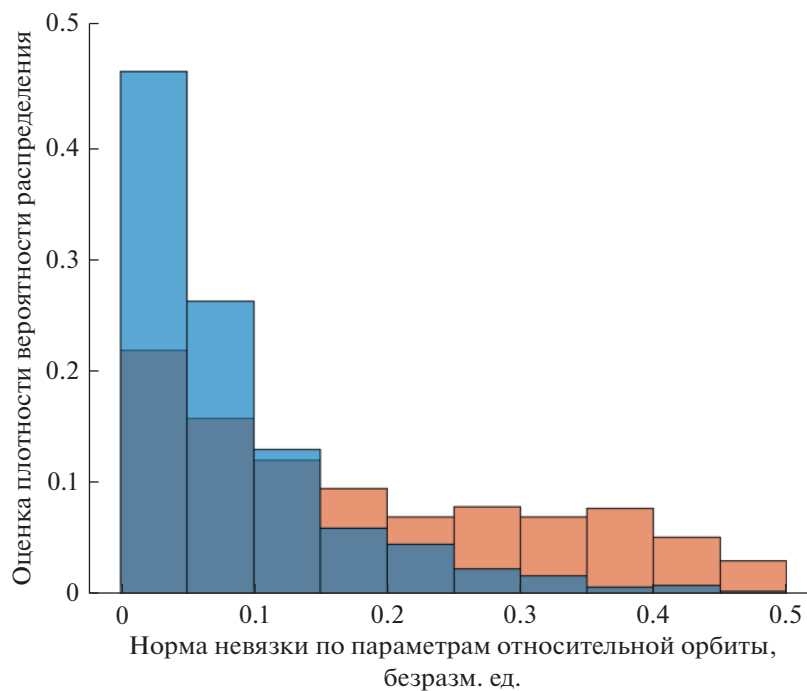


Рис. 4. Во время оптимизации число эпизодов равно 1, число шагов равно 5. Оценка распределения до и после оптимизации как на рис. 2.

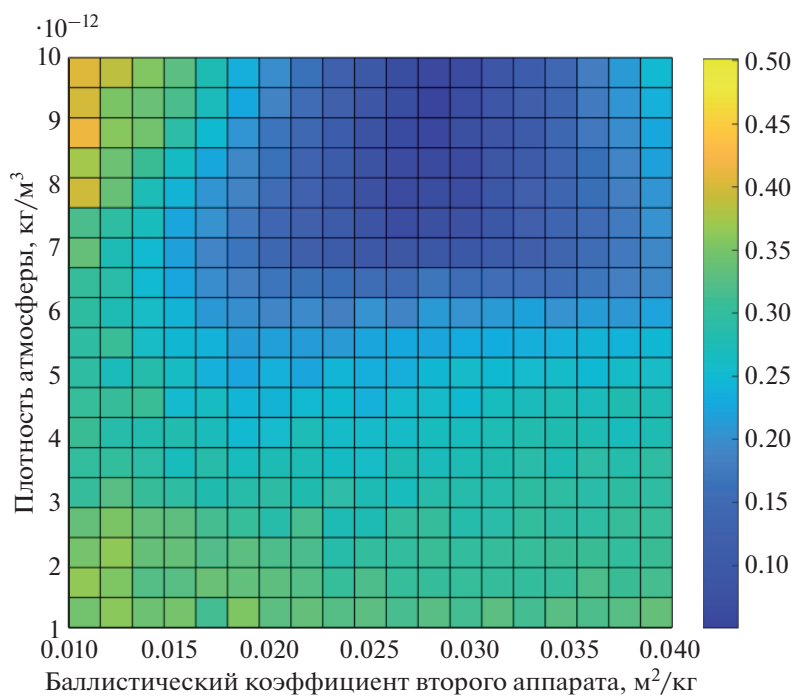


Рис. 5

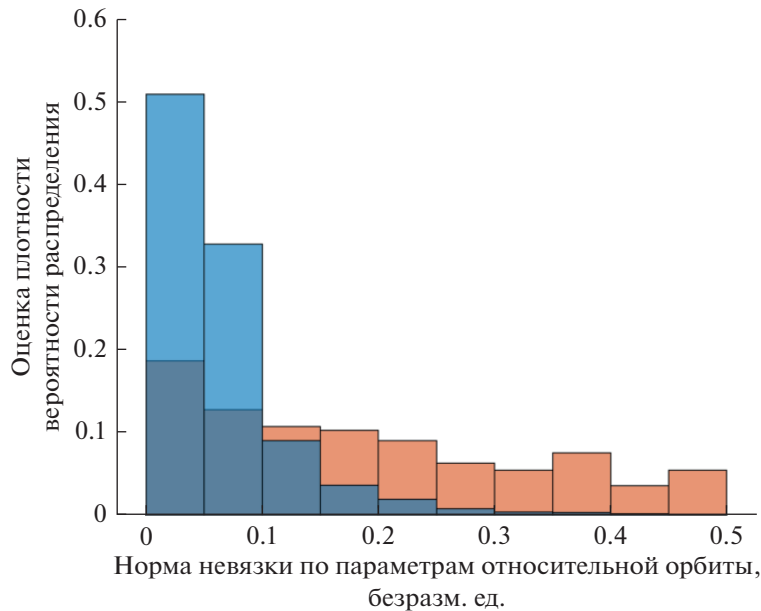


Рис. 6. Как на рис. 3.

формирований выборки для невязки Δ . Кроме того, численные эксперименты показали, что объем выборки из Δ должен быть в несколько раз больше, чем объем выборки, достаточный для адаптации только коэффициента β_2 . Например, если взять один эпизод и 5 шагов, то результат адаптации не будет удовлетворительным. Поэтому время полета от момента начала адаптации к параметрам до конца увеличивается в десятки раз. Из рис. 5 видно, что если оценка плотности атмосферы отклонена от истинного значения не более чем на $2 \cdot 10^{-12}$ кг/м³, то можно адаптировать только β_2 .

ЗАКЛЮЧЕНИЕ

Перечислим основные выводы, которые следуют из проделанной работы.

Задачу построения нейроуправления формацией двух спутников на низких околоземных орбитах лучше выстраивать в рамках методов обучения с учителем, нежели в рамках методов обучения с подкреплением.

Хотя в некоторых задачах идея введения управляющей и прогнозирующей нейронных сетей является естественной, в данном случае эта идея оказалась неэффективной, так как коррекция весов нейронных сетей дестабилизирует управление. В приведенной работе есть только управляющие нейронные сети, а задача адаптации нейроуправления может быть эффективно решена как задача подбора входных параметров управляющих сетей, а не как задача коррекции весов этих сетей.

Чтобы добиться высокой точности аппроксимации нейронными сетями функции управления, потребовалось использовать глубокие нейронные сети – с числом скрытых слоев не менее 3.

Средняя невязка попадания на номинальную относительную орбиту (проективную круговую) оказалась унимодальной функцией баллистического коэффициента неуправляемого аппарата и плотности атмосферы, рассматриваемых как вход в управляющие нейронные сети. Задачу адаптации (как оптимизационную задачу) можно решать неградиентными методами.

Удовлетворительные результаты адаптации получаются даже в случае, когда адаптируется только баллистический коэффициент второго аппарата: на сбор статистики о точности попадания на номинальную орбиту требуется около 8 ч, а вся процедура адаптации занимает от 2.5 до 5 дней. Лучше результаты адаптации получатся в случае, если адаптируются оба параметра – баллистический коэффициент и плотность атмосферы. Однако время для сбора статистики требуется увеличить в десятки раз.

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации (Соглашение №075-15-2020-808).

СПИСОК ЛИТЕРАТУРЫ

1. *Вьюгин В.В.* Математические основы машинного обучения и прогнозирования. М.: МЦНМО, 2014.
2. *Саттон Р.С., Барто Э.Г.* Обучение с подкреплением. М.: БИНОМ. Лаборатория знаний, 2011.
3. *KrishnaKumar K., Rickard S., Bartholomew S.* Adaptive neuro-control for spacecraft attitude control // Neuro-

- computing. 1995. V. 9. № 2. P. 131–148.
[https://doi.org/10.1016/0925-2312\(94\)00062-W](https://doi.org/10.1016/0925-2312(94)00062-W)
4. *Biggs J.D., Fournier H.* Neural-Network-Based Optimal Attitude Control Using Four Impulsive Thrusters // *J. Guidance, Control, and Dynamics*. 2020. V. 43. № 2. P. 299–309.
<https://doi.org/10.2514/1.G004226>
 5. *Wright W.A.* Stochastic tuning of a spacecraft controller using neural networks // *Engineering Applications of Artificial Intelligence*. 1995. V. 8. № 6. P. 651–656.
[https://doi.org/10.1016/0952-1976\(95\)00043-7](https://doi.org/10.1016/0952-1976(95)00043-7)
 6. *Domingo V., Fleck B., Poland A.I.* SOHO: The Solar and Heliospheric Observatory // *Space Science Reviews*. 1995. V. 72. № 1. P. 81–84.
<https://doi.org/10.1007/BF00768758>
 7. *Сорокин А.В., Широбоков М.Г.* Коррекция и прогнозирование орбитального движения космических аппаратов с помощью искусственных нейронных сетей. Препринты ИПМ им. М.В. Келдыша. 2018. № 198. С. 28.
<https://doi.org/10.20948/prepr-2018-198>
 8. *Сорокин А.В., Широбоков М.Г.* Разработка нейронных сетей для управления орбитальным движением космического аппарата с двигателем малой тяги. Препринты ИПМ им. М.В. Келдыша. 2018. № 269. С. 31.
<https://doi.org/10.20948/prepr-2018-269>
 9. *Cheng L., Wang Z., Jiang F. et al.* Real-Time Optimal Control for Spacecraft Orbit Transfer via Multiscale Deep Neural Networks // *IEEE Transactions on Aerospace and Electronic Systems*. 2018. V. 55. № 5. P. 2436–2450.
<https://doi.org/10.1109/TAES.2018.2889571>
 10. *Cheng L., Wang Z., Jiang F.* Real-time control for fuel-optimal Moon landing based on an interactive deep reinforcement learning algorithm // *Astrodynamics*. 2019. V. 3. № 4. P. 375–386.
<https://doi.org/10.1007/s42064-018-0052-2>
 11. *Gurfil P., Idan M., Kasdin N.J.* Adaptive Neural Control of Deep-Space Formation Flying // *J. Guidance, Control, and Dynamics*. 2003. V. 26. № 3. P. 491–501.
<https://doi.org/10.2514/2.5072>
 12. *Bae J., Kim Y.* Adaptive controller design for spacecraft formation flying using sliding mode controller and neural networks // *J. Franklin Institute*. 2012. V. 349. № 2. P. 578–603.
<https://doi.org/10.1016/j.jfranklin.2011.08.009>
 13. *Dachwald B.* Optimization of Interplanetary Solar Sailcraft Trajectories Using Evolutionary Neurocontrol // *J. Guidance, Control, and Dynamics*. 2004. V. 27. № 1. P. 66–72.
<https://doi.org/10.2514/1.9286>
 14. *Dachwald B.* Optimization of very-low-thrust trajectories using evolutionary neurocontrol // *Acta Astronautica*. 2005. V. 57. № 2–8. P. 175–185.
<https://doi.org/10.1016/j.actaastro.2005.03.004>
 15. *Dachwald B., Ohndorf A.* 1st ACT global trajectory optimisation competition: Results found at DLR // *Acta Astronautica*. 2007. V. 61. № 9. P. 742–752.
<https://doi.org/10.1016/j.actaastro.2007.03.011>
 16. *Carnelli I., Dachwald B., Vasile M.* Evolutionary Neurocontrol: A Novel Method for Low-Thrust Gravity-Assist Trajectory Optimization // *J. Guidance, Control, and Dynamics*. 2009. V. 32. № 2. P. 616–625.
<https://doi.org/10.2514/1.32633>
 17. *Izzo D., Märten M., Pan B.* A survey on artificial intelligence trends in spacecraft guidance dynamics and control // *Astrodynamics*. 2019. V. 3. № 4. P. 287–299.
<https://doi.org/10.1007/s42064-018-0053-6>
 18. *Izzo D., Simões L.F., de Croon G.C.H.E.* An evolutionary robotics approach for the distributed control of satellite formations // *Evolutionary Intelligence*. 2014. V. 7. P. 107–118.
<https://doi.org/10.1007/s12065-014-0111-9>
 19. *Mohan S., Saenz-Otero A., Nolet S. et al.* SPHERES flight operations testing and execution // *Acta Astronautica*. 2009. V. 65. № 7. P. 1121–1132.
<https://doi.org/10.1016/j.actaastro.2009.03.039>
 20. *Gaudet B., Furfaro R.* Adaptive pinpoint and fuel efficient mars landing using reinforcement learning // *IEEE/CAA J. Automatica Sinica*. 2014. V. 1. № 4. P. 397–411.
<https://doi.org/10.1109/JAS.2014.7004667>
 21. *Furfaro R., Wibben D.R., Gaudet B. et al.* Terminal Multiple Surface Sliding Guidance for Planetary Landing: Development, Tuning and Optimization via Reinforcement Learning // *J. Astronautical Sciences*. 2015. V. 62. № 1. P. 73–99.
<https://doi.org/10.1007/s40295-015-0045-1>
 22. *Jiang X., Li S., Furfaro R.* Integrated guidance for Mars entry and powered descent using reinforcement learning and pseudospectral method // *Acta Astronautica*. 2019. V. 163. P. 114–129.
<https://doi.org/10.1016/j.actaastro.2018.12.033>
 23. *Gaudet B., Linares R., Furfaro R.* Deep Reinforcement Learning for Six Degree-of-Freedom Planetary Landing // *Advances in Space Research*. 2020. V. 65. № 7. P. 1723–1741.
<https://doi.org/10.1016/j.asr.2019.12.030>
 24. *Gaudet B., Linares R., Furfaro R.* Adaptive guidance and integrated navigation with reinforcement meta-learning // *Acta Astronautica*. 2020. V. 169. P. 180–190.
<https://doi.org/10.1016/j.actaastro.2020.01.007>
 25. CIRA-2012: COSPAR international reference atmosphere – 2012. Models of the Earth’s Upper Atmosphere. 2012. P. 20–24. https://spaceweather.usu.edu/files/chapters_1_3.pdf
 26. *Hooke R., Jeeves T.A.* “Direct search” solution of numerical and statistical problems // *J. ACM*. 1961. V. 8. № 2. P. 212–229.
<https://doi.org/10.1145/321062.321069>