Адаптивное согласованное управление мультиагентными системами на основе нейронных сетей

В. Сюй[1], Ю. Юань[2]

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

[1]xuwenran@yandex.ru, [2]1449771962@qq.com

***Аннотация*. В данной статье предлагается децентрализованная архитектура адаптивного согласованного управления мультиагентными системами на основе дифференциальных уравнений в частных производных. Данная архитектура интегрирует нейронные сети на основе радиально-базисных функций и графовые нейронные сети, что позволяет реализовать масштабируемое распределенное обучение при одновременном обеспечении устойчивости замкнутой системы. С использованием метода функционалов Ляпунова строго доказана полуглобальная равномерная конечная ограниченность решений замкнутой системы. Результаты численного моделирования подтверждают эффективность предложенного подхода в обеспечении пространственно-временной синхронизации состояний.**

Ключевые слова: дифференциальные уравнения в частных производных, мультиагентные системы, нейронные сети на основе радиально-базисных функций, графовые нейронные сети

I. Введение

В условиях гетерогенной динамики и ограничений связи мультиагентные системы (МАС), моделируемые дифференциальными уравнениями в частных производных (ДУЧП), сталкиваются с серьезными вызовами при реализации согласованного управления. Традиционные централизованные архитектуры управления характеризуются вычислительными «узкими местами» и зависимостью от глобальной информации, что препятствует их эффективному масштабированию [1]. Применение методов адаптивного управления, таких как нейронные сети на основе радиально-базисных функций (РБФ-НС), представляется перспективным подходом к аппроксимации априорно неизвестной нелинейной динамики [2], однако вопросы их использования в системах МАС на основе ДУЧП при наличии структурных ограничений остаются недостаточно исследованными.

В последнее время графовые нейронные сети (ГНС), моделируя отношения взаимодействия между агентами в данных с графовой структурой, открыли новые пути реализации децентрализованного обучения [3]. Тем не менее, существующие подходы зачастую ориентированы на эвристические стратегии координации, пренебрегая требованиями строгого теоретического анализа устойчивости, необходимого для систем, описываемых ДУЧП. В задачах, требующих обеспечения глобальной сходимости на основе локальных взаимодействий, интеграция анализа устойчивости по Ляпунову с распределенным обучением все еще остается теоретическим пробелом в современной литературе [4].

Для решения указанных проблем в данной работе предлагается децентрализованная архитектура адаптивного управления на основе сопряжения ГНС и РБФ-НС, предназначенная для задач согласованного управления мультиагентными системами, моделируемыми ДУЧП. Предложенная структура позволяет агентам осуществлять совместную оптимизацию стратегий управления децентрализованным образом, опираясь исключительно на информацию о локальных ошибках и параметрах взаимодействия весов, передаваемую по топологии графа.

II. Децентрализованная архитектура управления мультиагентными системами на основе ДУЧП с сопряжением ГНС и РБФ-НС

A. Постановка задачи и основные допущения

Рассматривается мультиагентная система, состоящая из N нелинейных электромеханических объектов, пространственная область которых распределена в ограниченной непрерывной области $\Omega \subset \mathbb{R}^m (m \le 3)$ с гладкой границей $\partial\Omega$. Пусть $\xi \in \Omega$ — пространственная координата, а $t \in [0, \infty)$ — временная переменная. Пространственно-временная динамика i -го агента ($i = 1, \dots, N$) описывается следующим ДУЧП типа «реакция-диффузия»:

$$\frac{\partial x_i(\xi,t)}{\partial t} = \nabla_\xi \cdot (D_i(\xi)\nabla_\xi x_i(\xi,t)) + f_i(x_i(\xi,t),\xi) + u_i(\xi,t), (\xi,t) \in \Omega \times [0,\infty)$$

Здесь $x_i(\xi,t) \in \mathbb{R}^n$ — вектор состояния системы; $D_i(\xi) = d_i(\xi) I_n$ — строго положительно определенная матрица коэффициентов диффузии, удовлетворяющая условию $d_i(\xi) \ge d_{min} > 0$; $f_i(x_i, \xi) \in \mathbb{R}^n$ — непрерывная неизвестная пространственная нелинейная динамика; $u_i(\xi,t) \in \mathbb{R}^n$ — распределенный закон управления, приложенный внутри пространственной области.

Для обеспечения управляемости системы и строгости теоретических выводов вводятся следующие стандартные допущения:

Допущение 1 (Неориентированная связная топология связи): Предполагается, что физический граф топологии связи G нижнего уровня является неориентированным и связным, то есть элементы его матрицы смежности удовлетворяют условию $b_{ij} = b_{ji}$, а сам граф содержит остовное дерево.

Допущение 2 (Согласованность опорной траектории и граничных условий): Целевая опорная траектория $x_r(\xi,t) \in \mathbb{R}^n$ и ее частные производные всех порядков известны, ограничены и гладки. Также требуется выполнение однородного граничного условия Неймана $(D_i(\xi)\nabla_\xi x_r(\xi,t)) \cdot n = 0$, что гарантирует отсутствие дополнительных возмущающих потоков на границе в уравнениях динамики ошибки.

Допущение 3 Условие согласования и изоморфное отображение): предполагается, что неизвестная нелинейная составляющая f_i и вход управления u_i удовлетворяют условию согласования. Кроме того, все агенты имеют общее базовое неизвестное физическое отображение, то есть существуют общие идеальные параметры W^* для проведения параметризации.

Вектор локальной ошибки слежения для i -го агента определяется как $e_i(\xi,t) = x_i(\xi,t) - x_r(\xi,t)$.

B. *Синтез закона граничного согласованного управления и закона внутридоменного адаптивного управления с упреждением*

Для учета физического взаимодействия между агентами и обеспечения синхронизации их состояний система удовлетворяет неоднородным граничным условиям Неймана на границе области:

$$(D_i(\xi)\nabla_\xi x_i(\xi,t)) \cdot n = u_{i,b}(\xi,t),\ (\xi,t) \in \partial\Omega \times [0,\infty)$$

Пусть $P_i \in \mathbb{R}^{n\times n}$ — положительно определенная симметричная матрица, используемая в дальнейшем в функционале Ляпунова. Чтобы гарантировать строгое формирование отрицательно полуопределенной квадратичной формы граничного интегрального члена в процессе доказательства устойчивости, закон граничного согласованного управления с обратной связью выбирается согласованным с матрицей P_i в следующем виде:

$$u_{i,b}(\xi,t) = -P_i^{-1}K_b e_i(\xi,t) - cP_i^{-1}\sum\nolimits_{j\in N_i} a_{ij}(t)(e_i(\xi,t) \\ -e_j(\xi,t))$$

Здесь $K_b \in \mathbb{R}^{n\times n}$ — положительно определенная матрица коэффициентов усиления обратной связи, $c > 0$ — коэффициент согласования, а $a_{ij}(t)$ — весовые коэффициенты топологии графа.

В рамках адаптивной схемы управления внутри области вводится расширенный вектор входных сигналов $Z_i = [x_i^T, \xi^T]^T$. Согласно теореме об универсальной аппроксимации, на компактном множестве $Z \subset \mathbb{R}^{n+m}$ для любой непрерывной функции $f_i(\cdot)$ существует матрица идеальных весовых коэффициентов W^*, такая что:

$$f_i(x_i(\xi,t),\xi) = W^{*T}\phi_i(Z_i) + \epsilon_i(\xi,t)$$

здесь $\phi_i(\cdot)$ — ограниченный вектор радиально-базисных функций, $\|\phi_i(Z_i)\| \le \Phi$, а $\epsilon_i(\xi,t)$ — ограниченная ошибка аппроксимации, $\|\epsilon_i(\xi,t)\| \le \bar{\epsilon}$.

С учетом известной опорной траектории, закон управления внутри области $u_i(\xi,t)$ синтезируется как структура, включающая упреждающую компенсацию и робастное подавление:

$$u_i(\xi,t) = -P_i^{-1}K_i e_i(\xi,t) + u_{ff,i}(\xi,t) - \hat{W}_i^T(t)\phi_i(Z_i) \\ -\rho_i P_i^{-1}\mathrm{sat}\left(\frac{P_i e_i(\xi,t)}{\delta}\right)$$

где $K_i > 0$ — положительно определенная матрица обратной связи; $u_{ff,i}(\xi,t) = \frac{\partial x_r}{\partial t} - \nabla_\xi \cdot (D_i \nabla_\xi x_r)$ — точный член упреждающей компенсации, предназначенный для нейтрализации пространственно-временных возмущений опорной траектории x_r ;робастный насыщающий член строго согласован с вектором $P_i e_i$, при этом коэффициент усиления $\rho_i > \bar{\epsilon}$ выбирается таким образом, чтобы непосредственно компенсировать остаточную ошибку, возникающую из интегрального члена $\int_\Omega e_i^T P_i \epsilon_i d\xi$ в производной функционала Ляпунова.

C. *Топологическое взвешивание ГНС и закон адаптации в непрерывном времени*

Для обеспечения симметричной положительной полуопределенности матрицы Лапласа графа, исходные направленные признаки внимания $\alpha_{ij}(t)$, генерируемые модулем трансформера, подвергаются принудительной симметризации на основе топологии физических ребер связи b_{ij} :

$$a_{ij}(t) = b_{ij} \cdot \frac{1}{2}(\alpha_{ij}(t) + \alpha_{ji}(t))$$

Ошибка оценки весов нейронной сети определяется как $\tilde{W}_i(t) = \hat{W}_i(t) - W^*$. На основе данного определения предлагается следующий закон обновления параметров в непрерывном времени:

$$\dot{\hat{W}}_i(t) = \Gamma_i \int_\Omega \phi_i(Z_i)e_i^T(\xi,t)P_i d\xi - \Gamma_i\lambda\sum\nolimits_{j\in N_i} a_{ij}(t)(\hat{W}_i(t) \\ -\hat{W}_j(t))$$

Здесь $\Gamma_i > 0$ — матрица коэффициентов адаптации. Благодаря симметричности весов графа $a_{ij} = a_{ji} \ge 0$, второй член связи параметров при суммировании по всей системе для функционала Ляпунова естественным образом удовлетворяет следующему ключевому теоретико-графовому тождеству для квадратичной формы:

$$\sum\nolimits_{i=1}^N \sum\nolimits_{j\in N_i} a_{ij}\mathrm{tr}(\tilde{W}_i^T(\hat{W}_i - \hat{W}_j)) = \frac{1}{2}\sum\nolimits_{ij} a_{ij}\|\tilde{W}_i - \tilde{W}_j\|_F^2 \\ = \mathrm{tr}(\tilde{W}^T(L\otimes I)\tilde{W}) \ge 0$$

Данное свойство гарантирует, что закон адаптации обеспечивает теоретически доказуемую неположительно определенную диссипацию, что обусловливает согласованную сходимость всех локальных распределенных сетей.

III. Анализ устойчивости

Лемма 1. Для любой симметричной матрицы смежности графа $A = [a_{ij}]$, элементы которой удовлетворяют условиям $a_{ij} = a_{ji} \geq 0$ и $a_{ii} = 0$, для произвольной последовательности векторов v_i соответствующей размерности справедливо тождество: $\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij} v_i^T (v_i - v_j) = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij} \| v_i - v_j \|^2$.

Определение 1. Для вектора $z = [z_1, \ldots, z_n]^T \in \mathbb{R}^n$ функция насыщения по компонентам $sat(z)$ определяется как $sat(z) = [sat(z_1), \ldots, sat(z_n)]^T$, где $sat(z_k) = \text{sign}(z_k)\min(|z_k|, 1)$.

Теорема 1 (Полуглобальная равномерная конечная ограниченность). Рассмотрим пространственно-временную нелинейную мультиагентную систему, описываемую ДУЧП. Предположим, что матрица коэффициентов диффузии системы является пространственно-постоянной ($D_i(\xi) \equiv D_i > 0$) и выполняются допущения 1–3. При использовании синтезированного закона внутридоменного управления $u_i(\xi, t)$, закона граничного согласованного управления $u_{i,b}(\xi, t)$, а также закона адаптации параметров в непрерывном времени $\dot{\hat{W}}_i(t)$, если параметры управления удовлетворяют следующим условиям: Матрицы коэффициентов усиления обратной связи $K_i > 0$ и $K_b > 0$ являются симметричными и положительно определенными; Граф топологии связи G является неориентированным и связным, а матрица весовых коэффициентов удовлетворяет условиям $A(t) = A^T(t)$ и её элементы покомпонентно неотрицательны, что гарантирует симметричную положительную полуопределенность матрицы Лапласа графа $L(t) = D(t) - A(t)$; Робастный коэффициент усиления удовлетворяет условию $\rho_i \geq \bar{\epsilon} + \epsilon_0$, где $\epsilon_0 > 0$ — некоторая константа; Скорость адаптации и коэффициент связи удовлетворяют условиям $\Gamma_i > 0$ и $\lambda > 0$; Тогда все сигналы замкнутой системы, включая локальную пространственную ошибку слежения $e_i(\xi, t)$ и ошибку оценки весов нейронной сети $\tilde{W}_i(t)$, являются полуглобально равномерно конечно ограниченными.

Доказательство:

Сконструируем следующий составной функционал Ляпунова:

$$V(t) = \sum_{i=1}^{N} V_i(t) = \sum_{i=1}^{N} \left(\frac{1}{2} \int_{\Omega} e_i^T(\xi, t) P_i e_i(\xi, t) d\xi + \frac{1}{2} \text{tr}(\tilde{W}_i^T(t) \Gamma_i^{-1} \tilde{W}_i(t)) \right)$$

Производная $\dot{V}(t)$ вычисляется вдоль траекторий системы с учетом того, что $\dot{\tilde{W}}_i = \dot{\hat{W}}_i$.

Основные этапы доказательства:

Диффузионные и граничные члены: Применение первой формулы Грина к диффузионному члену $\nabla_{\xi} \cdot (D_i \nabla_{\xi} e_i)$ позволяет выделить граничные потоки. Подстановка закона граничного управления $u_{i,b}$ обеспечивает неположительность суммы граничных интегралов.

Нейросетевая адаптация: Благодаря структуре ГНС и симметризации весов a_{ij} , перекрестные члены ошибок весов $\tilde{W}_i$ образуют диссипативную квадратичную форму $\text{tr}(\tilde{W}^T (L \otimes I) \tilde{W}) \geq 0$, что гарантирует сходимость весов.

Робастное подавление: Влияние ошибки аппроксимации ϵ_i нейтрализуется робастным насыщающим членом с усилением $\rho_i > \bar{\epsilon}$.

Итоговая оценка устойчивости:

После объединения всех оценок и исключения отрицательно определенных членов, производная функционала удовлетворяет неравенству:

$$\dot{V}(t) \leq -\gamma V(t) + \Delta$$

где $\gamma = \dfrac{2\alpha}{\min_i \lambda_{\min}(P_i)}$ — коэффициент скорости сходимости, а Δ — суммарная остаточная погрешность. В силу леммы о сравнении справедливо выражение $lim\,sup_{t \to \infty} V(t) \leq \Delta / \gamma$, что означает достижение в замкнутой системе полуглобальной равномерной конечной ограниченности. Доказательство завершено.

IV. Результаты моделирования и анализ

Для подтверждения эффективности предложенной децентрализованной двухканальной архитектуры управления на основе сопряжения ГНС и РБФ-НС применительно к мультиагентным системам, описываемым ДУЧП, в данном разделе проведено численное моделирование в среде MATLAB.

На рис. 1 показан процесс динамической эволюции состояния репрезентативного агента 1 в непрерывной пространственно-временной области. Это подтверждает, что РБФ-НС эффективно восстанавливает неизвестное физическое отображение благодаря теореме об универсальной аппроксимации, а использование робастного насыщающего члена позволяет успешно подавлять остаточную ошибку аппроксимации $\epsilon_i(\xi, t)$:

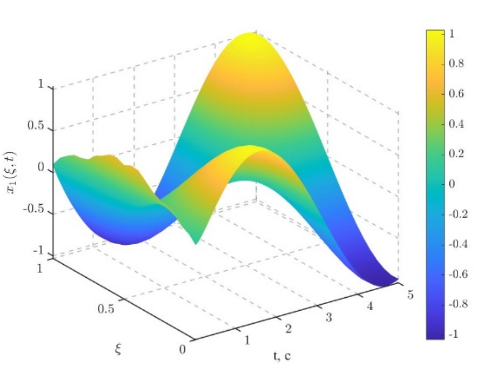

Рис. 1. Пространственно-временная эволюция состояния агента

На рис. 2 представлены кривые эволюции локальных норм L_2 -ошибок слежения $\| e_i(\cdot,t)\|_2$ пяти агентов при согласованном управлении; траектории ошибок всех пяти агентов практически совпадают. На протяжении всего периода моделирования ошибки слежения всех агентов остаются ограниченными в пределах крайне малой остаточной окрестности Δ , что свидетельствует о достижении высокоточной пространственно-временной синхронизации состояний.

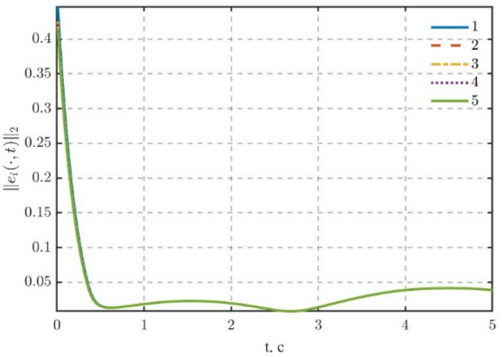

Рис. 2. Кривые ошибок слежения

Таким образом, результаты численного моделирования убедительно подтверждают справедливость Теоремы 1. Предложенная децентрализованная двухканальная архитектура управления в полностью автономном режиме успешно обеспечивает полуглобальную равномерную конечную ограниченность всех сигналов замкнутой мультиагентной системы, описываемой нелинейными ДУЧП типа «реакция-диффузия».

V. Заключение

В данной работе рассмотрена проблема согласованного управления мультиагентными системами, моделируемыми дифференциальными уравнениями в частных производных. Предложена децентрализованная адаптивная архитектура управления на основе сопряжения ГНС и РБФ-НС, гарантирующая полуглобальную равномерную конечную ограниченность решений замкнутой системы. Для компенсации неизвестной пространственной нелинейной динамики и учета ограничений топологии связи применены внутридоменное адаптивное управление с упреждением на базе РБФ-НС и граничное согласованное управление с обратной связью, согласованное с топологией графа. Теоретическая эффективность предложенных стратегий управления подтверждена строгим анализом с использованием композитных функционалов Ляпунова. Результаты численного моделирования дополнительно верифицируют работоспособность данной архитектуры в обеспечении пространственно-временной синхронизации состояний.

Список литературы

[1] De Gennaro M.C., Jadbabaie A. Decentralized control of connectivity for multi-agent systems. *Proceedings of the 45th IEEE Conference on Decision and Control*, 2006, pp. 3628-3633. DOI: 0.1109/CDC.2006.377041.

[2] Yang H., Liu J. An adaptive RBF neural network control method for a class of nonlinear systems. *IEEE/CAA Journal of Automatica Sinica*, 2018, vol. 5, no. 2, pp. 457-462. DOI: 10.1109/JAS.2017.7510820.

[3] Zhao T., Chen T., Zhang B. QMIX-GNN: A graph neural network-based heterogeneous multi-agent reinforcement learning model for improved collaboration and decision-making. *Applied Sciences*, 2025, vol. 15, no. 7, 3794. DOI: 10.3390/app15073794.

[4] Zhou P., Hu X., Zhu Z., Ma J. What is the most suitable Lyapunov function? *Chaos, Solitons & Fractals*, 2021, vol. 150, 111154. DOI: 10.1016/j.chaos.2021.111154.