مجله کنترل
ISSN (print) 2008-8345
ISSN (online) 2538-3752
جلد ۱۲، شماره ۳، پاییز ۱۳۹۷، صفحه ۲۸-۱۳

همزمانسازی بهینه برخط سیستم های چندعاملی غیر خطی با دینامیک های نامعلوم

فرزانه تاتاری [۱]، محمد باقر نقیبی سیستانی [۲]

۱ استادیار دانشکده مهندسی برق و کامپیوتر، گروه مهندسی برق، دانشگاه سمنان، ftatari@semnan.ac.ir

۲ دانشیار دانشکده مهندسی، گروه مهندسی برق، دانشگاه فردوسی مشهد، mb-naghibi@um.ac.ir

دریافت : ۱۳۹۶/۴/۱۱ ویرایش: ۱۳۹۶/۱۲/۲۹ پذیرش: ۱۳۹۷/۲/۱

چکیده: در این مقاله، الگوریتم بهینه توزیع شده تطبیقی برخط برای همزمانسازی عامل های غیرخطی یک سیستم چندعاملی با دینامیک های نامعلوم به عامل رهبر بر اساس تکنیک های برنامه ریزی پویای تقریبی و شناساگرهای شبکه های عصبی ارایه شده است. الگوریتم پیشنهاد شده به یادگیری حل برخط معادلات همیلتون-جاکوبی تزویج شده[۱] (CHJ) تحت دینامیک های نامعلوم پرداخته است. هر عامل جهت یادگیری سیاست بهینه محلی از ساختار عملگر-نقاد بهره برده و دینامیک نامعلوم هر عامل نیز با به کارگیری یک تقریبگر شبکه عصبی، تقریب زده شده است. شناسایی دینامیک های نامعلوم با استفاده از قانون تکرار تجربیات انجام شده است به طوری که از اطلاعات ثبت شده به همراه داده های لحظه ای برای انطباق وزن های شبکه عصبی شناساگر دینامیک عامل ها، استفاده شده است. در حالی که وزن های تقریبگرهای دینامیک و شبکه های عملگر-نقاد به صورت همزمان در حال انطباق هستند، کرانداری تمامی سیگنال های حلقه بسته توسط تئوری لیاپانوف تضمین شده است. در انتها صحت الگوریتم پیشنهاد شده با ذکر نتایج شبیه سازی، نشان داده شده است.

کلمات کلیدی: برنامه ریزی پویای تقریبی، تقریبگرهای عملگر-نقاد، سیستم های چندعاملی، کنترل بهینه توزیع شده، همزمانسازی.

Online optimal Synchronization of Nonlinear Multi-agent Systems under Unknown Dynamics

Farzaneh Tatari, Mohammad-B. Naghibi-S.

Abstract: In this paper an online optimal distributed algorithm is introduced for multi-agent systems synchronization under unknown dynamics based on approximate dynamic programming and neural networks. Every agent has employed an actor-critic structure to learn its distributed optimal policy and the unknown dynamics of every agent is identified by employing a neural network approximator. The unknown dynamics are identified based on the experience replay technique where the recorded data and current data are used to adopt the approximators weights. The introduced algorithm learns the solution of coupled Hamilton-Jacobi equations under unknown dynamics in an online fashion. While the weights of the identifiers and actor-critic approximators are being tuned, the boundedness of the closed loop system signals are assured using Lyapunov theory. The effectiveness of the proposed algorithm is shown through the simulation results.

Keywords: Actor-critic approximators; Approximate dynamic programming; Multi-agent systems; Optimal distributed control; Synchronization.

[۱] Coupled Hamilton-Jacobi

۱- مقدمه

سیستم های چندعاملی متشکل از چندین عامل همسان یا غیرهمسان می باشند که این عامل‌ها با همکاری یا رقابت با یکدیگر به حل یک مسئله می‌پردازند. استفاده از کنترل همکارانه جهت هماهنگ کردن رفتار عامل های خودمختار در محیط های با پهنای ارتباطی پایین امری چالش برانگیز است. تا کنون بیشتر مطالعات کنترل همکارانه بر روی رهیافت های متمرکز[۱]، تمرکز داشته اند که همه عامل ها نیاز دارند تا به طور پیوسته با یک عامل مرکزی در ارتباط باشند. استفاده از این روش باعث ایجاد ترافیک ارتباطی سنگین، تاخیر و از دست رفتن اطلاعات می شود. بعلاوه عامل مرکزی نیازمند دارا بودن منابع محاسباتی کافی است تا داده ها را پردازش کرده و سیگنال پاسخ را تولید کند. مشکلات روش های متمرکز نیاز به رهیافت توزیع شده[۲] را ایجاب می کند. در کنترل توزیع شده سیستم های چند عاملی، به طراحی قانون کنترلی ساده برای هر عامل با استفاده از اطلاعات همسایگانش پرداخته می شود به طوری که سیستم چندعاملی توزیع شده بتواند به رفتار جمعی (اجماع) دست پیدا کند.

در مسئله اجماع یا همزمانسازی پیرو-رهبر با طراحی کنترل مناسب برای عامل ها، متغیر حالت همه عامل ها به حالت یک عامل کنترلی یا رهبر [۳-۱] همگرا می شود. حل مسئله اجماع پیرو-رهبر[۳] (ردیابی همکارانه) در کنترل همکارانه سیستم های چندعاملی خطی [۴،۵] و غیرخطی [۶،۷] تحت گرافهای ارتباطی با توپولوژی ثابت و متغیر با زمان، مورد بررسی قرار گرفته اند که در تمام آنها دینامیک سیستم کاملا معین بوده و روش های به کار رفته برون خط می باشند و معیار بهینگی برای تعیین کنترل مناسب در نظر گرفته نشده است. بخشی از مسائل کنترل همکارانه سیستم های چندعاملی را کنترل بهینه همکارانه این سیستم ها تشکیل می دهد. اکثر تکنیک های به کار رفته برای حل مسئله اجماع در کنترل بهینه همکارانه سیستم های چندعاملی، برون خط می باشند [۸،۹] و در اکثر این روش ها سعی می شود به نحوی از حل معادلات پیچیده هامیلتون جاکوبی تزویج شده[۴] (CHJ) اجتناب شود.

نظریه بازی های دیفرانسیلی شامل ترکیبی از مشخصه های تئوری بازی [۱۰] و تئوری کنترل بهینه می باشد [۱۱]. در نتیجه می توان گفت، بازی های دیفرانسیلی تعمیم مسائل کنترل بهینه هستند که بیش از یک بازیکن یا عامل در بازی نقش دارند [۱۲،۱۳]. تئوری بازی دیفرانسیلی راهکار مناسبی برای مدل کردن مسائل تصمیم گیری چند عاملی برای سیستم های دیفرانسیلی در تعامل می باشد [۱۴]. در این بازی ها هر عامل مستقل از دیگران، به بهینه سازی شاخص عملکرد خود می پردازد تا سیاست بهینه خود را تعیین کند و نتیجه آن حل تعادل نش[۵] بازی می باشد. پیدا کردن راه حل نش بازی، وابسته به حل معادلات پیوستهCHJ می باشد، حل این معادلات دیفرانسیل جزئی غیرخطی بسیار مشکل بوده و یا حتی در مواردی فاقد حل تحلیلی همه جایی می باشند. حل معادلات همیلتون-جاکوبی تزویج شده به صورت برون خط انجام شده است [۱۴-۱۶]. اما مزایای استفاده از سیاست های کنترلی بهینه برخط این است که شرایطی را برای سیستم های چند عاملی فراهم می کند تا عامل‌ها بتوانند اهداف و معیارهای بهینگی خود را به صورت برخط تغییر دهند و بتوانند سیاست متناسب با موقعیت جدید را محاسبه نمایند.

یادگیری تقویتی[۶] شاخه ای از یادگیری ماشین است که در آن، یک یا چندین عامل به صورت زمان-حقیقی برای رسیدن به یک هدف خاص که همان استراتژی بهینه است، با محیطی که ممکن است برای عامل ها شناخته شده نباشد، تعامل می نمایند[۱۷]. این عامل ها استراتژی بهینه را بر اساس تجربیاتی که در جهت بیشینه کردن مجموعه پاداش خود کسب می کنند، یاد می گیرند. در یادگیری تقویتی با استفاده از مقدار اسکالر شاخص عملکرد، که سیگنال تقویت یا پاداش نامیده می شود، فرآیند آموزش به عامل ها در محیط های غیرقطعی و پیچیده انجام می شود.

می توان از روشهای یادگیری تقویتی برای حل برخط مسائل کنترل بهینه همکارانه توزیع شده و بازی های گرافی دیفرانسیلی[۷] استفاده نمود. در بازی های گرافی دیفرانسیلی، دینامیک خطا و اندیس عملکرد هر عامل بستگی به عامل های همسایه اش دارد و یا به عبارت دیگر بستگی به ساختار گراف ارتباطی شبکه دارد. در بازی های گرافی دیفرانسیلی هدف یافتن مجموعه ای از سیاست های کنترلی قابل قبول برای سیستم چندعاملی است که منجر به همزمان سازی عامل ها، تضمین پایداری سیستم حلقه بسته و حداقل شدن تابع هزینه هر عامل در جهت رسیدن به تعادل نش، شود. برای پیاده سازی روش های یادگیری تقویتی برخط نیاز به استفاده از تئوری تقریب[۸]داریم، به این منظور از برنامه ریزی پویای تقریبی[۹](ADP) استفاده می شود [۱۸]. برنامه ریزی پویای تقریبی کلاسی از الگوریتم هاست که راه حل های برخط تقریبی را برای مسائل کنترل بهینه فراهم می آورد. ADP با استفاده از نمایش تقریبی تابع ارزش و کنترلگر و به کارگیری اصل بهینگی بلمن که اساس برنامه ریزی پویا می باشد، روشی را برای تعلیم برخط ساختارهای تقریبگر ارزش و کنترلگر فراهم می آورد [۱۹].

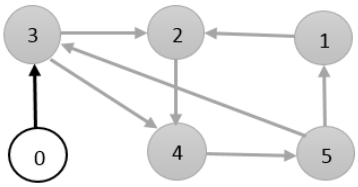
از روش تکرار سیاست و ساختارهای عملگر-نقاد برای کنترل بهینه برخط بازی های دیفرانسیلی مجموع غیر صفر استفاده شده است[۲۰-۲۲]. به منظور اجتناب از اطلاعات متمرکز که در این مقاله ها استفاده شده است،

[۱] Centralized
[۲] Distributed
[۳] Leader-follower consensus
[۴] Coupled Hamilton-Jacobi
[۵] Nash Equilibrium Solution
[۶] Reinforcement Learning
[۷] Differential graphical games
[۸] Approximation theory
[۹] Approximate Dynamic programming

در [۲۳] بازی های گرافی دیفرانسیلی برای سیستم های خطی با دینامیک معین معرفی شده و با به کارگیری ساختارهای عملگر-نقاد به حل برخط معادلات CHJ پرداخته شده است. مرجع [۲۴] یک روش تکرار سیاست توسعه یافته برخط را برای حل تقریبی برخط معادلات CHJ برای بازی های گرافی دیفرانسیلی خطی زمان پیوسته معرفی کرد که در آن تنها دینامیک داخلی برای هر عامل نامعلوم می باشد. تاتاری و همکارانش در [۲۵]، برای حل برخط مسئله همزمانسازی عامل ها در بازی های گرافی دیفرانسیلی خطی با دینامیک کاملا نامعلوم، از ساختارهای عملگر-نقاد و تقریبگرهای محلی به ترتیب برای تقریب سیاست های بهینه بازیکنان و شناسایی دینامیک عامل ها به صورت زمان حقیقی استفاده کرده اند. بازی های گرافی دیفرانسیلی غیرخطی برای سیستم های زمان پیوسته در [۲۶] معرفی شدند که الگوریتم ارایه شده در این مرجع تضمین پایداری حلقه بسته و همگرایی سیاست ها به نقطه تعادل نش را تضمین می کند. در [۲۷،۲۸]، همزمانسازی عامل ها به رهبر به صورت بازی گرافی دیفرانسیلی با استفاده از یادگیری تقویتی برون-سیاست[۱] که نیازمند به مدل سیستم نمی باشد، انجام شده است. بعلاوه در این مراجع، عامل ها دارای دینامیک خطی داخلی یکسان می باشند و برای به روزرسانی وزن های شبکه های عصبی نقاد و عملگر از قوانین استاندارد گرادیان نزولی استفاده شده است. در حالی که در مقاله پیش رو، عامل ها دارای دینامیک غیرخطی و کاملا متمایز می باشند و برای حل مسئله از روش های یادگیری تقویتی بر-سیاست[۲] استفاده شده است. که با حل مسئله همزمانسازی، دینامیک نامعین هر یک از عامل ها نیز شناسایی می شود. در این مقاله نیز برای به روزرسانی وزن های شبکه های عصبی نقاد و عملگر برای هر یک از عامل ها به ترتیب از قوانین استاندارد گرادیان نزولی و قوانین غیر استاندارد استفاده می شود.

با توجه به مطالعات انجام شده تا کنون مطالعه ای بر روی حل بهینه برخط بازی های گرافی دیفرانسیلی غیرخطی تحت دینامیک نامعلوم انجام نشده است. بنابراین نوآوری اصلی این مقاله معرفی یک الگوریتم یادگیری توزیع شده بهینه برای حل بازی های گرافی دیفرانسیلی غیرخطی با دینامیک نامعلوم است. در اینجا هر بازیکن از تقریبگرهای عملگر و نقاد به ترتیب برای یادگیری ارزش بهینه و سیاست کنترلی بهینه استفاده می کند. به طور همزمان با پیاده سازی ساختار عملگر-نقاد، دینامیک غیرخطی نامعلوم هر عامل به صورت برخط با استفاده از شبکه های عصبی شناساگر که از تکنیک تکرار تجربیات بهره می برند شناسایی می شود. نشان داده می شود که الگوریتم پیشنهاد شده به حل معادلات CHJ همگرا می شود و سیاست های کنترلی بهینه به دست آمده به صورت تقریبی به حل تعادل نش بازی همگرا می شوند. کرانداری همه سیگنال های حلقه بسته نیز با استفاده از نظریه لیاپانوف تضمین می گردد.

مفاهیم مربوط به گرافها: توپولوژی تبادل اطلاعات بین N عامل، توسط گراف $Gr(V,\Sigma)$ توصیف می شود. $V=\{1,2,...,N\}$ مجموعه گره های گراف است که نماینده N عامل می باشد، $\Sigma\subseteq V\times V$ مجموعه شاخه های گراف و $(i,j)\in\Sigma$ به معنی وجود یک شاخه از گره i به گره j می باشد. توپولوژی یک گراف معمولا توسط ماتریس همسایگی آن $E=\left[e_{ij}\right]\in\Re^{N\times N}$ نمایش داده می شود به طوری که اگر $(j,i)\in\Sigma$ آنگاه $e_{ij}=1$ و در غیر این صورت $e_{ij}=0$ می باشد. $N_i=\{j:(j,i)\in\Sigma\}$ مجموعه همسایگان گره i است، به عبارت دیگر مجموعه گره ها با شاخه هایی است که به گره i وارد می شوند. $i_{N_i}=\{j:(i,j)\in\Sigma\}$ نیز نشان دهنده مجموعه ای از عامل ها هستند که عامل i در همسایگی آنها می باشد. $d_i=\sum_{j\in N_i}e_{ij}$ درجه-واردشونده[۳] گره i می باشد که مجموع عناصر سطر i ام E می باشد. مسیر، دنباله ای از گره های به هم متصل در یک گراف است و یک گراف را متصل گویند اگر مسیری بین هر دو گره دلخواه آن وجود داشته باشد. معمولا گره رهبر توسط اندیس صفر نشان داده می شود و اطلاعات از رهبر به عامل هایی که رهبر در همسایگی آنهاست، فرستاده می شوند. شکل ۱ نمونه ای از گراف ارتباطی یک سیستم چندعاملی را نشان میدهد.



شکل ۱: گراف ارتباطی یک سیستم چندعاملی

۱- ۲- تشریح مسئله

دینامیک N عامل که بر روی گراف ارتباطی Gr با یکدیگر در تعامل هستند را به صورت زیر در نظر بگیرید.

$$\dot{x}_i=f_i(x_i)+g_i(x_i)\,u_i\ ,\quad i=1,...,N. \qquad (۱)$$

در معادلات فوق $x_i\in\Re^n$ حالت قابل اندازه گیری عامل i ام، $f_i(x_i)\in\Re^n$ دینامیک داخلی سیستم، $g_i(x_i)\in\Re^{n\times m}$ دینامیک ورودی سیستم و $u_i\in\Re^m$ ورودی کنترلی عامل i ام می باشد. توابع $f_i(x_i)$ و $i=1,...,N, g_i(x_i)$ لیپشیتز محلی هستند و بر روی یک مجموعه فشرده تعریف شده اند. بعلاوه سیستم (۱)

[۱] Off-policy
[۲] On-policy
[۳] In-degree

پایدارپذیر است. فرض می شود که دینامیک های $f_i(x_i)$ و $g_i(x_i)$ برای $i=1,...,N$ نامعلوم هستند. دینامیک عامل رهبر نیز به صورت زیر بوده که در آن $x_0 \in \Re^n$ حالت رهبر می باشد

$$\dot{x}_0 = f_0(x_0) \qquad (۲)$$

هدف مسئله، همزمان سازی حالت همه عامل ها به حالت رهبر ضمن شناسایی دینامیک های نامعلوم می باشد و طراحی پروتکل های کنترل محلی تنها با استفاده از اطلاعات عامل های همسایه می باشد. خطای ردیابی محلی $\delta_i \in \Re^n$ به صورت زیر تعریف می گردد:

$$\delta_i = \sum_{j\in N_i} e_{ij}(x_i - x_j) + e_{i0}(x_i - x_0) \qquad (۳)$$

با مشتق گیری از (۱) و استفاده از (۲) و (۳)، دینامیک خطای ردیابی محلی به صورت زیر به دست می آید:

$$\dot{\delta}_i = \sum_{j\in N_i} e_{ij}(\dot{x}_i - \dot{x}_j) + e_{i0}(\dot{x}_i - \dot{x}_0) =$$
$$\sum_{j\in N_i} e_{ij}(f(x_i) - f(x_j)) + e_{i0}(f(x_i) - f(x_0)) \qquad (۴)$$
$$+(d_i + e_{i0})g_i(x_i)u_i - \sum_{j\in N_i} e_{ij}g_j(x_j)u_j$$

به منظور رسیدن به اجماع با استفاده از قوانین کنترلی محلی بهینه، تابع ارزش متناظر با گره i به صورت زیر می باشد

$$V_i(\delta_i(t)) = \frac{1}{2}\int_t^\infty (Q_i(\delta_i) + u_i^T R_{ii} u_i + \sum_{j\in N_i} u_j^T R_{ij} u_j)\, d\tau \qquad (۵)$$

که در آن $Q_i(\delta_i) > 0$ معمولا تابعی غیر خطی می باشد و ماتریس های وزن $R_{ii} > 0, R_{ij} > 0$ ، ثابت و متقارن هستند. سیاست های نقطه تعادل نش $u_i^* \in \Omega_i, i \in N$ برای بازی گرافی دیفرانسیلی با N بازیکن، مجموعه استراتژی های $\{u_1^*, u_2^*, ..., u_N^*\}$ هستند که $V_i^* = V_i(\delta_i(0), u_i^*, u_{N_i}^*) \le V_i(\delta_i(0), u_i, u_{N_i}^*)$ به ازای همه $u_i^* \in \Omega,\ i \in N$ برقرار باشد [۱۴و۲۳].

هدف بازی گرافی دیفرانسیلی یافتن مقادیر ارزش بهینه، $\forall \delta_i(t)$ به صورت زیر می باشد

$$V_i(\delta_i(t)) = \min_{u_i} \frac{1}{2}\int_t^\infty \left(Q_i(\delta_i) + u_i^T R_{ii} u_i + \sum_{j\in N_i} u_j^T R_{ij} u_j \right) d\tau \qquad (۶)$$

وقتی V_i محدود می شود، با مشتق گرفتن از (۵)، معادله لیاپانفی به شکل معادله بلمن زیر با شرط اولیه $V_i(0) = 0$ به دست می آید.

$$\nabla V_i^T [\sum_{j\in N_i} e_{ij}(f(x_i) - f(x_j)) + e_{i0}(f(x_i)$$
$$-f(x_0)) + (d_i + e_{i0})g_i(x_i)u_i - \sum_{j\in N_i} e_{ij}g_j(x_j)u_j] \qquad (۷)$$
$$+\frac{1}{2}Q_i(\delta_i) + \frac{1}{2}u_i^T R_{ii} u_i + \frac{1}{2}\sum_{j\in N_i} u_j^T R_{ij} u_j = 0$$

که در آن $\nabla V_i = \partial V_i / \partial \delta_i \in \Re^n$. حل رابطه (۷) روشی دیگر برای ارزیابی انتگرال نامحدود (۶) برای پیدا کردن ارزش مربوط به سیاست های فیدبکی، موجود می باشد. برای دینامیک (۴) و شاخص عملکرد (۵)، همیلتونین مربوطه به صورت زیر به دست می آید.

$$H_i(\delta_i, \nabla V_i, u_i, u_{N_i}) \equiv \nabla V_i^T [\sum_{j\in N_i} e_{ij}(f(x_i)$$
$$-f(x_j)) + e_{i0}(f(x_i) - f(x_0)) + (d_i + e_{i0}) \times$$
$$g_i(x_i)u_i - \sum_{j\in N_i} e_{ij}g_j(x_j)u_j] + \frac{1}{2}Q_i(\delta_i) \qquad (۸)$$
$$+\frac{1}{2}u_i^T R_{ii} u_i + \frac{1}{2}\sum_{j\in N_i} u_j^T R_{ij} u_j$$

با استفاده از اصل بهینگی بلمن، کنترل فیدبک بهینه u_i^* ، به صورت زیر به دست می آید.

$$u_i^* = u_i^*(V_i^*) = -(d_i + e_{i0})R_{ii}^{-1} g_i^T(x_i)\nabla V_i^* \qquad (۹)$$

با جاگذاری سیاست های کنترلی (۹) در (۷)، معادلات CHJ پیوسته برای سیستم های غیرخطی به صورت زیر به دست می آید

$$\nabla V_i^{*T} [\sum_{j\in N_i} e_{ij}(f(x_i) - f(x_j)) + e_{i0}(f(x_i)$$
$$-f(x_0)) - (d_i + e_{i0})^2 g_i(x_i)R_{ii}^{-1} g_i^T(x_i)\nabla V_i^*$$
$$-\sum_{j\in N_i} e_{ij}(d_j + e_{j0})g_j(x_j)R_{jj}^{-1} g_j^T(x_j)\nabla V_j^*] \qquad (۱۰)$$
$$+\frac{1}{2}Q_i(\delta_i) + \frac{1}{2}(d_i + e_{i0})^2 \nabla V_i^{*T} g_i(x_i)R_{ii}^{-1} \times$$
$$g_i^T(x_i)\nabla V_i^* + \frac{1}{2}\sum_{j\in N_i} (d_j + e_{j0})^2 \nabla V_j^{*T} \times$$
$$g_j(x_j)R_{jj}^{-1}R_{ij}R_{jj}^{-1} g_j^T(x_j)\nabla V_j^* = 0$$

متناظر با هر گره یک معادله CHJ وجود دارد، بنابراین حل بازی گرافی دیفرانسیلی N نفره با سیستم غیرخطی، مستلزم حل N معادله دیفرانسیل CHJ ذکر شده می باشد. از آنجایی که حل معادلات دیفرانسیل CHJ (۱۰) در حالت کلی بسیار مشکل می باشد و نامعلوم بودن دینامیک های $f_i(x_i)$ و $g_i(x_i)$ $i=1,...,N$ برای هر عامل i این پیچیدگی را افزایش می دهد، با استفاده از روش های تقریبی به شناسایی و حل این بازی ها می پردازیم.

ملاحظه ۱- بنا به وجود جملات تزویج شده در معادلات CHJ، روش هایی که تا کنون برای حل تقریبی برخط معادلات هامیلتون جاکوبی بلمن[۱] HJB معرفی شده [۲۹-۳۱] را نمی توان مستقیما برای حل معادلات CHJ بازی های گرافی دیفرانسیلی توسعه داد و حل برخط تقریبی این معادلات از پیچیدگی بیشتری برخوردار است. بعلاوه حل تقریبی برخط معادلات CHJ برای بازی های گرافی دیفرانسیلی غیرخطی، به علت وجود دینامیک های غیرخطی و نامعلوم در سیستم چند عاملی، متفاوت و پیچیده تر از حل موجود در [۲۶،۲۵] خواهد بود.

برای ادامه نیاز به ذکر تعریف زیر داریم.

تعریف ۱ (تحریک پایا[۲] (PE))- سیگنال برداری کراندار $\bar{X}_i(t),\ i=1,...,N$ روی بازه $[t,t+T_i]$ تحریک پایا می باشد [۳۱] اگر وجود داشته باشند $T_i>0, \gamma_i>0$ و $\gamma_{i+N}>0$ به طوری که برای هر t ،

$$\gamma_i I \le \int_t^{t+T_i} \bar{X}_i(\tau)\bar{X}_i^{\ T}(\tau)d\tau \le \gamma_{i+N} I \ .$$

۳- شناسایی سیستم چند عاملی غیرخطی با استفاده از شبکه های عصبی

با الهام از [۳۰]، شناساگرهای شبکه عصبی برای مدل کردن دینامیک نامعلوم عامل ها به کار گرفته می شوند. با فرض اینکه دینامیک $f_i(x_i)$ و $g_i(x_i)$ عامل ها پیوسته بوده و روی یک مجموعه فشرده تعریف شده اند از N شناساگر شبکه عصبی برای تقریب آنها به صورت زیر استفاده می شود

$$\begin{aligned} &f_i(x_i)=\theta_i^{*}\xi_i(x_i)+\varepsilon_{f_i}\ ,\\ &g_i(x_i)=\psi_i^{*}\varsigma_i(x_i)+\varepsilon_{g_i}\ ,\ i=1,...,N \end{aligned} \qquad (۱۱)$$

که در آن $\theta_i^{*}\in\Re^{n\times k_{\theta_i}}$ و $\psi_i^{*}\in\Re^{n\times k_{\psi i}}$ وزن های نامعلوم شبکه های عصبی هستند و $\xi_i\in\Re^{k_{\theta_i}}$ و $\varsigma_i\in\Re^{k_{\psi_i}\times m}$ توابع پایه شبکه عصبی هستند. ε_{f_i} و ε_{g_i} خطای تقریب شبکه های عصبی هستند.

فرض ۱- برای تقریبگرهای شبکه های عصبی به صورت استاندارد فرض می شود که خطای شبکه عصبی و گرادیان آن روی مجموعه فشرده Ω کراندار هستند و توابع فعالیت شبکه عصبی و گرادیان آنها نیز کراندار هستند.

ملاحظه ۲- فرض ۱ از جمله فرض های استاندارد در ادبیات شبکه های عصبی می باشد [۲۹،۳۱،۳۲]. به عنوان مثال قسمت اول فرض ۱ برای توابع فعالیت سیگموید، تانژانت هیپربولیک و دیگر توابع فعالیت استاندارد شبکه عصبی برقرار می باشد. مراجع [۲۶،۲۳،۲۰] به صورت مشابه این فرض را لحاظ کرده اند.

سیستم (۴) می تواند به صورت زیر نوشته شود

$$\dot{x}_i=\varphi_{f_ig_i}^{*}z(x_i,u_i)+\varepsilon_{f_ig_i},\ i=1,...,N \qquad (۱۲)$$

که در آن $\varphi_{A_iB_i}^{*}=[\theta_i^{*}\quad \psi_i^{*}]$ و $z(x_i,u_i)=[\xi_i^{T}\quad u_i^{T}\varsigma_i^{T}]^{T}$ بردار رگسور است و $\varepsilon_{f_ig_i}=\varepsilon_{f_i}+\varepsilon_{g_i}$. با استفاده از فرض ۱ داریم $\left\|\varepsilon_{f_ig_i}\right\|\le\bar{\varepsilon}_{f_ig_i}$ که $\bar{\varepsilon}_{f_ig_i}=\bar{\varepsilon}_{f_i}+\bar{\varepsilon}_{g_i}$ ، $\left\|\varepsilon_{f_i}\right\|\le\bar{\varepsilon}_{f_i}$ و $\left\|\varepsilon_{g_i}\right\|\le\bar{\varepsilon}_{g_i}$. بنابراین $\varepsilon_{f_ig_i}$ کراندار است. دینامیک (۱۲) می تواند به صورت زیر نوشته شود $i=1,...,N$

$$\dot{x}_i=-Ax_i+\varphi_{f_ig_i}^{*}z(x_i,u_i)+Ax_i+\varepsilon_{f_ig_i}, \qquad (۱۳)$$

که در آن $A=aI_{n\times n},\ a>0$.

لم ۱- [۳۰] حل سیستم (۱۳) به صورت زیر قابل بازنویسی است

$$x_i=\varphi_{f_ig_i}^{\ *}h_i(x_i)+al_i(x_i)+\varepsilon_{x_i} \qquad (۱۴)$$

$$\begin{aligned} &\dot{h}_i(x_i)=-ah_i(x_i)+z(x_i,u_i)\ ,\\ &h_i(0)=0\ ,\ h_i(x_i)\in\Re^{k_{\theta_i}+k_{\psi_i}}\\ &\dot{l}_i(x_i)=-Al_i(x_i)+x_i\ ,\ l_i(0)=0 \end{aligned} \qquad (۱۵)$$

که در آن $h_i(x_i)=\int_0^t e^{-a(t-\tau)}z(x_i(\tau),u_i(\tau))d\tau$ معادل بردار رگسور فیلتر شده $z(x_i,u_i)$ می باشد، $x_i(0)$ حالت اولیه سیستم و $l_i(x_i)=\int_0^t e^{-A(t-\tau)}x_i(\tau)d\tau$ ، $\varepsilon_{x_i}=e^{-At}x_i(0)+\int_0^t e^{-A(t-\tau)}\varepsilon_{f_ig_i}\,d\tau$ می باشد.

■

جهت ایجاد قانون تطبیق برای هر عامل i، ابتدا رابطه (۱۴) را به سیگنال نرمالایز کننده $n_{s_i}=1+h_i^{T}h_i+l_i^{T}l_i$ تقسیم می کنیم و داریم

$$\bar{x}_i=\varphi_{f_ig_i}^{\ *}\bar{h}_i(x_i)+a\bar{l}_i(x_i)+\bar{\varepsilon}_{x_i} \qquad (۱۶)$$

[۱] Hamilton-Jacobi Bellman

[۲] Persistently Exciting

که در آن $\bar{x}_i = \frac{x_i}{n_{s_i}}, \bar{h}_i = \frac{h_i}{n_{s_i}}, \bar{l}_i = \frac{l_i}{n_{s_i}}, \bar{\varepsilon}_{x_i} = \frac{\varepsilon_{x_i}}{n_{s_i}}$ به ترتیب شکل نرمالایز شده $x_i, h_i, l_i, \varepsilon_{x_i}$ می باشند. بر اساس لم ۱ و (۱۶)، تخمین حالت عامل i ام به صورت زیر می باشد.

$$\hat{\bar{x}}_i = \hat{\varphi}_{f_i g_i} \bar{h}_i(x_i) + a\bar{l}_i(x_i), \mathrm{i} = 1,\ldots,\mathrm{N} \quad (۱۷)$$

که در آن $\hat{\varphi}_{f_i g_i} = [\hat{\theta}_i, \hat{\psi}_i] \in \Re^{n\times(k_{\theta_i}+k_{\psi_i})}$ مقادیر تخمین زده شده ماتریس $\varphi_{f_i g_i}^*$ مربوط به عامل i در زمان t می باشد. خطای تخمین حالت برای هر عامل i به صورت زیر به دست می‌آید

$$e_i(t) = \hat{\bar{x}}_i - \bar{x}_i = \tilde{\varphi}_{f_i g_i}(t)\bar{h}_i(x_i(t)) - \bar{\varepsilon}_{x_i}, \quad \tilde{\varphi}_{f_i g_i}(t) = \hat{\varphi}_{f_i g_i}(t) - \varphi_{f_i g_i}^*(t) \quad (۱۸)$$

که در آن $\tilde{\varphi}_{f_i g_i}(t) = [\tilde{\theta}_i \quad \tilde{\psi}_i]$ خطای تخمین پارامتر برای عامل i در زمان t می باشد به طوری که $\tilde{\theta}_i = \hat{\theta}_i - \theta_i^*$, $\tilde{\psi}_i = \hat{\psi}_i - \psi_i^*, i = 1,\ldots,N$. برای به دست آوردن قانون آموزش وزن های شبکه های عصبی شناساگر از تکنیک تکرار تجربیات [۳۰] که مبتنی بر استفاده از داده های ثبت شده قبلی به همراه داده های فعلی است، استفاده می شود. به این منظور، بردار داده های ذخیره شده در زمان های قبلی $t_1,\ldots,t_{p_i}$ ، برای هر عامل $i = 1,\ldots,\mathrm{N}, i$ به صورت زیر تعریف می شود.

$$Z_i = [\bar{h}_i(x_i(t_1)),\ldots,\bar{h}_i(x_i(t_{p_i}))] \quad (۱۹)$$

به این ترتیب با استفاده از (۱۹)، قانون آموزش وزن های شبکه عصبی شناساگر عامل i به صورت زیر به دست می آید

$$\dot{\hat{\varphi}}_{f_i g_i}(t) = -\Gamma_i \bar{h}_i(x_i(t))e_i^T(t) - \Gamma_i \sum_{k=1}^{p_i} \bar{h}_i(x_i(t_k))e_i^T(t_k),\ i = 1,\ldots,N \quad (۲۰)$$

به طوری که p_i اندازه حافظه یا اندازه بردار داده های ذخیره شده قبلی برای هر عامل i و $i = 1,\ldots,N,\ \Gamma_i > 0$ ماتریس نرخ یادگیری مثبت معین می باشد و t نشان دهنده زمان فعلی است. بعلاوه $e_i(t_k)$ خطای تخمین حالت عامل i برای k امین داده ذخیره شده در زمان t_k می باشد که به صورت زیر تعریف می شود

$$e_i(t_k) = \hat{\bar{x}}_i(t,t_k) - \bar{x}_i(t_k) = \tilde{\varphi}_{f_i g_i}(t)\bar{h}_i(x_i(t_k)) - \bar{\varepsilon}_{x_i}(t_k), k = 1,\ldots,p_i. \quad (۲۱)$$

توجه شود که در قانون تنظیم (۲۰) جمله آخر به خطاهای تخمین زده شده قبلی و رگرسور فیلترشده نرمالایز شده عامل i بستگی دارد. با استفاده از قانون تنظیم (۲۰) در صورتی که شرط رتبه Z_i برقرار باشد، کران خطای وزن های شناسایی و خطای تخمین حالت عامل i به ناحیه کوچکی نزدیک صفر میل می کند.

به منظور برآورده کردن شرط PE، باید تعداد داده های مستقل خطی موجود در Z_i برابر با بعد پایه عدم قطعیت $h_i(x_i)$ در رابطه (۱۵) باشد که به آن شرط رتبه Z_i گوییم.

ملاحظه ۳- با استفاده از قانون آموزش گرادیان همزمان در رابطه (۲۰)، شرط محافظه کارانه PE حذف و به جای آن همگرایی پارامترهای شبکه شناساگر با برقراری شرط رتبه Z_i تضمین می شود.

تعریف ۲ (پایداری کراندار نهایی یکنواخت[۱](UUB))- سیگنال وابسته به زمان $\vartheta(t)$ کراندار نهایی یکنواخت گفته می شود اگر مجموعه فشرده $\Omega \subset R^n$ وجود داشته باشد به طوری که برای همه $\vartheta(0) \in \Omega$ یک کران β و زمان $T(\beta, \vartheta(0))$ وجود داشته باشد به طوری که برای همه $t \geq t_0 + T$ داشته باشیم $\|\vartheta(t)\| \leq \beta$ [۲۹].

قضیه ۱- سیستم (۱۲) را در نظر بگیرید. قانون تنظیم وزن های شبکه عصبی عامل i را مطابق با (۲۰) در نظر بگیرید. آنگاه در صورت برقراری شرط رتبه Z_i ، خطای تقریب مدل (خطای تخمین وزن های شناساگر) کراندار نهایی یکنواخت ($\left\|\tilde{\varphi}_{f_i g_i}\right\| \leq b_{\tilde{\varphi}_{f_i g_i}}$) می باشد.

اثبات قضیه ۱- مشابه مرجع [۳۰] انجام می شود.

■

با استفاده از (۱۲)، سیستم (۴) می تواند به صورت زیر نوشته شود

$$\dot{\delta}_i = \sum_{j\in N_i} e_{ij}(\theta_i^*\xi_i(x_i) - \theta_j^*\xi_j(x_j)) + e_{i0}\times (\theta_i^*\xi_i(x_i) - \theta_0^*\xi_0(x_0)) + (d_i + e_{i0})\psi_i^*\varsigma_i(x_i)u_i - \sum_{j\in N_i} e_{ij}\psi_j^*\varsigma_j(x_j)u_j + \varepsilon_{T_i}, \quad (۲۲)$$

$$\varepsilon_{T_i} = \sum_{j\in N_i} e_{ij}(\varepsilon_{f_i} - \varepsilon_{f_j}) + e_{i0}(\varepsilon_{f_i} - \varepsilon_{f_0}) + (d_i + e_{i0})\varepsilon_{g_i}u_i - \sum_{j\in N_i} e_{ij}\varepsilon_{g_j}u_j, i = 1,\ldots,N,\ j \in N_i$$

خطای تقریب دینامیک خطای محلی عامل می باشد. با استفاده از فرض ۱ و $\left\|\varepsilon_{f_i}\right\| \leq \bar{\varepsilon}_{f_i}$ و $\left\|\varepsilon_{g_i}\right\| \leq \bar{\varepsilon}_{g_i}$ داریم $\left\|\varepsilon_{T_i}\right\| \leq \bar{\varepsilon}_{T_i}$.

سیستم (۲۲) می تواند به صورت زیر نوشته شود $i = 1,\ldots,N$

[۱] Uniformly Ultimately Bounded

$$\dot{\delta}_i = \varphi_i^{*} z_i(x_i, x_j, u_i, u_{N_i}) + \varepsilon_{T_i}, \qquad (۲۳)$$

که در آن بردار رگرسور

$$z_i(x_i, x_j, u_i, u_{N_i}) = [z_{e_{ij}\xi_i}{}^T, z_{-e_{ij}\xi_j}{}^T, e_{i0}\xi_i{}^T, -e_{i0}\xi_0{}^T, (d_i + e_{i0})(\varsigma_i u_i)^T, z_{-e_{ij}\varsigma_j u_j}{}^T]^T \in \Re^{d^i}$$

$$z_{e_{ij}\xi_i} = \{e_{ij}\xi_i \mid j \in N_i\},\ z_{-e_{ij}\xi_j} = \{-e_{ij}\xi_j \mid j \in N_i\},$$

$$z_{-e_{ij}\varsigma_j u_j} = \{-e_{ij}\varsigma_j u_j \mid j \in N_i\},$$

$$Card(z_{-e_{ij}\varsigma_j u_j}) = Card(z_{-e_{ij}\xi_j}) = Card(z_{e_{ij}\xi_i}) = \|N_i\|,$$

$$d^i = \sum_{j \in N_i} (k_{\theta_i} + k_{\theta_j}) + k_{\theta_i} + k_{\theta_0} + k_{\psi_i} + \sum_{j \in N_i} k_{\psi_j}$$

می باشد و $\varphi_i^{*} = [\varphi_{\theta_i}{}^{*}, \varphi_{\theta_j}{}^{*}, \theta_i^{*}, \theta_0^{*}, \psi_i^{*}, \varphi_{\psi_j}{}^{*}] \in \Re^{n \times d^i}$ ماتریس وزن های نامعلوم می باشد، $\varphi_{\theta_i}{}^{*} = [\theta_i^{*}, ..., \theta_i^{*}]$، $\varphi_{\theta_j}{}^{*} = \{\theta_j^{*} \mid j \in N_i\}$، $\varphi_{\psi_j}{}^{*} = \{\psi_j^{*} \mid j \in N_i\}$ که در آنها $Card(\varphi_{\theta_i}{}^{*}) = Card(\varphi_{\theta_j}{}^{*}) = Card(\varphi_{\psi_j}{}^{*}) = \|N_i\|$ و $\|N_i\|$ نشان دهنده تعداد عامل های همسایه بازیکن i می باشد.

می توان تقریب (۲۳) را به صورت زیر نوشت $i = 1, ..., N$

$$\dot{\delta}_i = \hat{\varphi}_i z_i(x_i, x_j, u_i, u_{N_i}) + \varepsilon_{T_i}, \qquad (۲۴)$$

مقادیر $\hat{\varphi}_i = [\hat{\varphi}_{\theta_i}, \hat{\varphi}_{\theta_j}, \hat{\theta}_i, \hat{\theta}_0, \hat{\psi}_i, \hat{\varphi}_{\psi_j}] \in \Re^{n \times d^i}$ تخمین زده شده برای ماتریس φ_i^{*} مربوط به عامل i می باشد که در آن $\hat{\varphi}_{\theta_j} = \{\hat{\theta}_j \mid j \in N_i\}$، $\hat{\varphi}_{\psi_j} = \{\hat{\psi}_j \mid j \in N_i\}$، $\hat{\varphi}_{\theta_i} = [\hat{\theta}_i, ..., \hat{\theta}_i]$. بدین ترتیب $\tilde{\varphi}_i(t)$ به صورت زیر تعریف می شود و از قضیه ۱ نتیجه می شود که کراندار نهایی یکنواخت می باشد (یعنی $\|\tilde{\varphi}_i\| \le b_{\tilde{\varphi}_i}$).

$$\tilde{\varphi}_i(t) = \hat{\varphi}_i(t) - \varphi_i^{*}(t). \qquad (۲۵)$$

٤- حل برخط همزمانسازی سیستم چندعاملی غیر خطی با دینامیک های نامعلوم

در این بخش، الگوریتم عملگر-نقاد برای یادگیری حل بهینه برخط بازی های گرافی دیفرانسیلی غیرخطی زمان پیوسته دارای دینامیک نامعلوم، ارایه می گردد.

٤-۱- تقریب توابع هزینه عامل های غیرخطی توسط شبکه های عصبی نقاد

بر اساس قانون تقریب مرتبه بالای وایرشتراس [۳۲-۳۴]، مجموعه های پایه مستقل و کامل $\sigma_i(\delta_i) : \Omega \to \Re^{K^i}, i = 1, ..., N$ به طوری که $\sigma_i(0) = 0$ ، $\nabla\sigma_i(0) = 0$ و وزن های ثابت شبکه عصبی نقاد $W_i \in \Re^{K^i}$ ، $i = 1, ..., N$ وجود دارند به گونه ای که حل V_i برای (۷) و $\nabla V_i = \frac{\partial V_i}{\partial \delta_i}$ بر روی مجموعه فشرده Ω به صورت زیر تقریب زده می شوند، $i = 1, ..., N$.

$$V_i = W_i^T \sigma_i(\delta_i) + \omega_i(\delta_i), \qquad (۲۶)$$

$$\nabla V_i = \nabla\sigma_i{}^T W_i + \nabla\omega_i, \qquad (۲۷)$$

که در آن $\sigma_i(\delta_i) \in \Re^{K^i}$ بردارهای پایه توابع فعالیت شبکه عصبی هستند، $K^i, i = 1, ..., N$ تعداد نرون های لایه مخفی، $\omega_i(\delta_i)$ خطای تقریب شبکه های عصبی می باشد، $\nabla V_i = \frac{\partial V_i}{\partial \delta_i}$ ، $\nabla\sigma_i = \frac{\partial \sigma_i}{\partial \delta_i}$ و $\nabla\omega_i = \frac{\partial \omega_i}{\partial \delta_i}$. اگر تعداد نرون های لایه مخفی $K^i \to \infty$ ، خطای تقریب به طور یکنواخت $\omega_i \to 0$ و $\nabla\omega_i \to 0$ و $\sigma_i(\delta_i)$ تشکیل یک مجموعه پایه مستقل می دهد [۳۲،۳۳]. بر مبنای فرض ۱، روی مجموعه فشرده Ω، داریم $\|\omega_i\| \le b_{\omega_i}, \forall i$ ، $\|\nabla\omega_i\| \le b_{\nabla\omega_i}$ ، $\|\sigma_i\| \le b_{\sigma_i}$ و $\|\nabla\sigma_i\| \le b_{\nabla\sigma_i}$.

با به کارگیری شبکه های عصبی تقریبگر توابع ارزش، که شبکه های عصبی نقاد نامیده می شوند (معادلات (۲۶) و (۲۷))، و سیاست های فیدبکی u_i و u_{N_i} ، هامیلتونین (۸) به صورت زیر به دست می آید.

$$H_i(\delta_i, W_i, u_i, u_{N_i}) = W_i^T \nabla\sigma_i[\varphi_i^{*} z_i(\delta_i, \delta_j, u_i, u_{N_i}) + \varepsilon_{T_i}] + \frac{1}{2}Q_i(\delta_i) + \frac{1}{2}u_i{}^T R_{ii} u_i + \frac{1}{2}\sum_{j \in N_i} u_j{}^T R_{ij} u_j = e_{B_i} \qquad (۲۸)$$

$$e_{B_i} = -(\nabla\omega_i)^T \times [\varphi_i^{*} z_i(x_i, x_j, u_i, u_{N_i}) + \varepsilon_{T_i}],\ i = 1, ..., N \qquad (۲۹)$$

e_{B_i} خطای باقی مانده ناشی از تقریب شبکه های عصبی می باشند. تحت فرض ۱، این خطاهای باقی مانده روی مجموعه فشرده Ω کراندار هستند، $Sup_{x \in \Omega}\|e_{B_i}\| \le \bar{e}_i$, $i = 1, ..., N$.

فرض ۲- مشابه [۲۰] فرض می شود که برای یک مجموعه فشرده $\Omega \subset \Re^n$ و $i = 1, ..., N$:

الف) $f_i(x_i) \le b_f \|x_i\|$. ب) $g_i(x_i)$ توسط یک ثابت مشخص کراندار است $\|g_i(x_i)\| \le b_{g_i}$. ج) وزن های شبکه های عصبی نقاد توسط ثابت های مشخص کراندار هستند $\|W_i\| < W_{i\max}$.

در واقع وزن های ایده آل $W_i, i = 1, ..., N$ و φ_i^{*} ، نامعلوم می باشند و باید به صورت زمان حقیقی تقریب زده شوند. بنابراین،

خروجی شبکه های عصبی نقاد $\hat{V}_i(\delta_i)$ و معادلات بلمن تقریبی می توانند به ترتیب به صورت زیر نوشته شوند.

$$\hat{V}_i = \hat{W}_i^T \sigma_i(\delta_i) \quad (۳۰)$$

$$e_{H_i} = \hat{W}_i^T \nabla\sigma_i(\hat{\varphi}_i\left[z_i(x_i, x_j, u_i, u_{N_i})\right]^T) + \frac{1}{2}Q_i(\delta_i) + \frac{1}{2}u_i^T R_{ii} u_i + \frac{1}{2}\sum_{j\in N_i} u_j^T R_{ij} u_j \quad (۳۱)$$

که در آن $\hat{\varphi}_i = [\hat{\varphi}_{\theta_i}, \hat{\varphi}_{\theta_j}, \hat{\theta}_i, \hat{\theta}_0, \hat{\psi}_i, \hat{\varphi}_{\psi_j}]$ و $\hat{W}_i$ به ترتیب مقادیر تخمین زده شده از φ_i^* و W_i برای $i = 1,\ldots,N$ می باشند. اکنون مسئله پیدا کردن تابع ارزش برای هر عامل به تنظیم پارامترهای شبکه های عصبی نقاد $\hat{W}_i$ تبدیل شده است به گونه ای که خطای اختلاف زمانی e_{H_i} حداقل شود. تابع هدف زیر را در نظر بگیرید.

$$E_i = \frac{1}{2}e_{H_i}^T e_{H_i} \quad (۳۲)$$

قوانین تنظیم وزن های شبکه عصبی نقاد با استفاده از الگوریتم گرادیان نزولی نرمالایز شده به صورت زیر به دست می آید

$$\dot{\hat{W}}_i = -\alpha_i \frac{\partial E_i}{\partial \hat{W}_i} = -\alpha_i e_{H_i}\frac{\partial e_{H_i}}{\partial \hat{W}_i} = -\alpha_i \frac{\hat{B}_i}{(1+\hat{B}_i^T\hat{B}_i)^2} e_{H_i} = -\alpha_i \frac{\bar{B}_i}{\mathrm{m}_{s_i}} e_{H_i} \quad (۳۳)$$

که در آن

$$m_{s_i} = 1+\hat{B}_i^T\hat{B}_i,\ \bar{B}_i = \frac{\hat{B}_i}{1+\hat{B}_i^T\hat{B}_i},$$
$$\hat{B}_i = \nabla\sigma_i(\hat{\varphi}_i\left[z_i(x_i, x_j, u_i, u_{N_i})\right]),$$

$\alpha_i > 0,\ i = 1,\ldots,N$ نرخ یادگیری و $(1+\hat{B}_i^T\hat{B}_i)^2$ برای نرمالایزسازی به کار گرفته شده است.

لم ۲- $(u_i, u_{N_i}), \forall i$ را به عنوان یک مجموعه سیاست فیدبکی کراندار قابل قبول در نظر بگیرید و (۳۳) را به عنوان قانون تنظیم شبکه های عصبی نقاد و (۲۰) را برای تنظیم وزن های شبکه عصبی شناساگر در نظر بگیرید و فرض کنید که $\bar{B}_i$ تحریک پایا باشد. آنگاه برای خطاهای تقریب کراندار، خطای پارامترهای نقاد به صورت نمایی به مجموعه باقی مانده زیر همگرا می شود.

$$\eta_{i_1} e^{-\eta_{i_2} t} + \frac{\alpha_i}{m_{s_i}\eta_{i_2}} b_{\nabla\sigma_i}\|W_i\| \times \left(\|z_i(x_i, x_j, u_i, u_j)_{j\in N_i}\| b_{\tilde{\varphi}_i} + \bar{\varepsilon}_{T_i}\right) + \frac{\alpha_i}{m_{s_i}\eta_{i_2}}\bar{e}_i$$

اثبات لم ۲- با توجه به محدودیت صفحات حذف شده است.

۲-۴- تقریب سیاست های کنترلی عامل های غیرخطی با استفاده از شبکه ها ی عصبی عملگر

سیاست های کنترلی ایده آل با در نظر گرفتن $V_i(\delta_i) = W_i^T\sigma_i, i = 1,\ldots,N$ ، به صورت زیر به دست می آید

$$u_i = -(\mathrm{d}_i + \mathrm{e}_{i0})R_{ii}^{-1}(\psi_i^*\varsigma_i + \varepsilon_{g_i})^T\nabla\sigma_i^T W_i. \quad (۳۴)$$

با به کارگیری شبکه های عصبی عملگر، تخمین سیاست های کنترلی بهینه به صورت زیر نوشته می شود

$$\hat{u}_i = -(\mathrm{d}_i + \mathrm{e}_{i0})R_{ii}^{-1}(\hat{\psi}_i\varsigma_i)^T\nabla\sigma_i^T\hat{W}_{i+N} \quad (۳۵)$$

که $\hat{W}_{i+N}$ مقدار تخمین زده شده از وزن ایده آل شبکه عصبی W_i و $\hat{\psi}_i$ مقادیر تخمین زده از وزن های ایده آل $\psi_i^*, i = 1,\ldots,N$ می باشد. خطاهای تخمین شبکه های عصبی نقاد و عملگر به ترتیب در روابط (۳۶) و (۳۷) تعریف شده است

$$\tilde{W}_i = W_i - \hat{W}_i \quad (۳۶)$$

$$\tilde{W}_{i+N} = W_i - \hat{W}_{i+N} \quad (۳۷)$$

با استفاده از (۳۳) و توجه به این که ورودی های کنترلی از رابطه (۳۵) داده می‌شوند، قانون تنظیم شبکه های عصبی نقاد برای هر عامل i ، به صورت زیر به دست می آید

$$\dot{\hat{W}}_i = -\alpha_i \frac{\hat{B}_{i+N}}{(1+\hat{B}_{i+N}^T\hat{B}_{i+N})^2}[\hat{B}_{i+N}^T\hat{W}_i + \frac{1}{2}Q_i(\delta_i) + \frac{1}{2}(d_i + e_{i0})^2\hat{W}_{i+N}^T\hat{D}_i\hat{W}_{i+N} + \frac{1}{2}\sum_{j\in N_i}(d_j + e_{j0})^2\hat{W}_{j+N}^T\hat{E}_j\hat{W}_{j+N}] \quad (۳۸)$$

به این ترتیب، برای تضمین پایداری سیستم حلقه بسته و همگرایی سیاست های کنترلی به تعادل نش، قوانین تنظیم شبکه های عصبی عملگر به فرم غیر استاندارد به صورت زیر انتخاب می شوند

$$\dot{\hat{W}}_{i+N} = -\alpha_{i+N}\{(S_i\hat{W}_{i+N} - F_i\hat{W}_i) - \hat{D}_i\hat{W}_{i+N}\times \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}}\hat{W}_i - \sum_{j\in i_{N_i}}\hat{E}_j\hat{W}_{i+N}\frac{\bar{B}_{j+N}^T}{2m_{s_j}}\hat{W}_j\} \quad (۳۹)$$

که در آن $F_i \in \Re^{K_i\times K_i}$ و $S_i \in \Re^{K_i\times K_i}$ ماتریس های مثبت معین قطری هستند و

$$\hat{E}_j = \frac{\partial \sigma_j}{\partial \delta_j}(\hat{\psi}_j \varsigma_j) R_{jj}^{-T} R_{ij} R_{jj}^{-1} (\hat{\psi}_j \varsigma_j)^T \frac{\partial \sigma_j}{\partial \delta_j}^T,$$

$$\hat{D}_i = \nabla \sigma_i (\hat{\psi}_i \varsigma_i) R_{ii}^{-T} (\hat{\psi}_i \varsigma_i)^T \nabla \sigma_i^T,$$

$$\hat{B}_{i+N} = \nabla \sigma_i (\hat{\varphi}_i \left[z_i(x_i, x_j, u_{i+N}, u_{-(i+N)}) \right]),$$

$$u_{-(i+N)} = \{u_{j+N} \mid \mathrm{j} \in \mathrm{N}_i\},$$

$$m_{s_{i+N}} = 1 + \hat{B}_{i+N}^T \hat{B}_{i+N},\ i_{N_i} = \{ j : (i,j) \in \Sigma \},$$

$$\bar{B}_{i+N} = \frac{\hat{B}_{i+N}}{1 + \hat{B}_{i+N}^T \hat{B}_{i+N}},\ i = 1, \ldots, \mathrm{N}.$$

لم ۳ (نامساوی یونگ)- برای هر دو بردار x و y و هر $\varepsilon > 0$ ، رابطه زیر برقرار می باشد [۳۵]

$$x^T y \le \varepsilon \frac{\|x\|^2}{2} + \frac{\|y\|^2}{2\varepsilon}$$

■

۳-۴ تحلیل همگرایی و پایداری

در ادامه قضایایی که پایداری حلقه بسته و همگرایی برخط به پاسخ تقریبی بهینه بازی چندعاملی گرافی غیرخطی با دینامیک های نامعلوم را تضمین می کند مطرح شده است.

قضیه ۲ (پایداری سیستم حلقه بسته و همگرایی شبکه‌های عصبی عملگر-نقاد)- سیستم دینامیکی (۲۲) با وزن های نامشخص θ_i^* ، ψ_i^* و $\psi_j^*|_{j\in N_i}, i = 1, \ldots, \mathrm{N}$ و تعاریف بازی گرافی دیفرانسیلی چندعاملی را در نظر بگیرید. در حالی که شرط رتبه Z_i برقرار است، وزن های شبکه عصبی شناساگر از رابطه (۲۰) به روزرسانی می شود. ورودی های کنترلی از رابطه (۳۵) داده می شوند و قانون تنظیم شبکه عصبی نقاد برای عامل i با رابطه (۳۸) ارایه می شود و قانون تنظیم شبکه عصبی عملگر متناظر با عامل i از رابطه (۳۹) داده می شود. $\bar{B}_{i+N}, \forall i$ تحریک پایا می باشد و فرض های ۱ و ۲ برقرار می باشد. آنگاه با به کارگیری تعداد کافی نرون برای شبکه های عصبی، حالت های خطای سیستم حلقه بسته $\delta_i(t)$ و خطای شبکه عصبی نقاد $\tilde{W}_i, i = 1, \ldots, N$ و خطای شبکه عصبی عملگر $\tilde{W}_{i+N}, i = 1, \ldots, N$ کراندار نهایی یکنواخت هستند.

اثبات قضیه ۲- پیش از انجام اثبات لازم به ذکر است که در مواقع مورد نیاز، $\theta_i^* \xi_i + \varepsilon_{f_i}$ و $\psi_i^* \varsigma_i + \varepsilon_{g_i}$ به ترتیب به عنوان شکل های معادل توابع f_i و g_i استفاده می شوند. تابع لیاپانوف و مشتق این تابع را به ترتیب به صورت زیر در نظر بگیرید

$$L(t) = \sum_{i=1}^{N} \{ V_i(t) + \overbrace{\frac{1}{2}\tilde{W}_i^T \alpha_i^{-1} \tilde{W}_i}^{L_i(t)} + \overbrace{\frac{1}{2}\tilde{W}_{i+N}^T \alpha_{i+N}^{-1} \tilde{W}_{i+N}}^{L_{i+N}(t)} \} \quad (۴۰)$$

$$\dot{L}(t) = \sum_{i=1}^{N} \{ \dot{V}_i(t) \overbrace{-\tilde{W}_i^T \alpha_i^{-1} \dot{\hat{W}}_i}^{\dot{L}_i(t)} \overbrace{-\tilde{W}_{i+N}^T \alpha_{i+N}^{-1} \dot{\hat{W}}_{i+N}}^{\dot{L}_{i+N}(t)} \} \quad (۴۱)$$

اولین جمله (۴۱) با استفاده از (۲۸)، (۱۸) و (۳۴)-(۳۷) و انجام پاره ای از محاسبات به صورت زیر به دست می آید

$$\begin{aligned}
\sum_{i=1}^{N} \dot{V}_i(t) = \sum_{i=1}^{N} \{ & \dot{\bar{L}}_{V_i} - \frac{1}{2} Q_i(\delta_i) + W_i^T \frac{\partial \sigma_i}{\partial \delta_i} \times \\
& \{ -(\mathrm{d}_i + \mathrm{e}_{i0})^2 (\psi_i^* \varsigma_i + \varepsilon_{g_i}) R_{ii}^{-1} \times \\
& ((\tilde{\psi}_i + \psi_i^*) \varsigma_i)^T \nabla \sigma_i^T W_i + (\mathrm{d}_i + \mathrm{e}_{i0})^2 \times \\
& (\psi_i^* \varsigma_i + \varepsilon_{g_i}) R_{ii}^{-1} ((\tilde{\psi}_i + \psi_i^*) \varsigma_i)^T \nabla \sigma_i^T \tilde{W}_{i+N} \\
& + \sum_{j \in N_i} e_{ij} ((\mathrm{d}_j + \mathrm{e}_{j0}) (\psi_j^* \varsigma_j + \varepsilon_{g_j}) R_{jj}^{-1} \times \\
& ((\tilde{\psi}_j + \psi_j^*) \varsigma_j)^T \nabla \sigma_j^T W_j \\
& - (\mathrm{d}_j + \mathrm{e}_{j0}) (\psi_j^* \varsigma_j + \varepsilon_{g_j}) R_{jj}^{-1} \times \\
& ((\tilde{\psi}_j + \psi_j^*) \varsigma_j)^T \nabla \sigma_j^T \tilde{W}_{j+N} \} \} + \sum_{i=1}^{N} \omega_{i0}
\end{aligned} \quad (۴۲)$$

که در آن

$$\begin{aligned}
\sum_{i=1}^{N} \omega_{i0} = \sum_{i=1}^{N} \{ & \frac{\partial \omega_i}{\partial \delta_i} [\sum_{j \in N_i} e_{ij} (f_i(x_i) - f_j(x_j)) \\
& + e_{i0} (f_i(x_i) - f_0(x_0)) - (d_i + e_{i0})^2 g_i(x_i) \times \\
& R_{ii}^{-1} (\hat{\psi}_i \varsigma_i)^T \frac{\partial \sigma_i}{\partial \delta_i}^T \hat{W}_{i+N} + \sum_{j \in N_i} e_{ij} (d_j + e_{j0}) \times \\
& g_j(x_j) R_{jj}^{-1} (\hat{\psi}_j \varsigma_j)^T \frac{\partial \sigma_j}{\partial \delta_j}^T \hat{W}_{j+N}] \},
\end{aligned}$$

$$\begin{aligned}
\dot{\bar{L}}_{V_i} = & -W_i^T \frac{\partial \sigma_i}{\partial \delta_i} [-(\mathrm{d}_i + \mathrm{e}_{i0})^2 (\psi_i^* \varsigma_i + \varepsilon_{g_i}) \times \\
& R_{ii}^{-1} (\psi_i^* \varsigma_i + \varepsilon_{g_i})^T \nabla \sigma_i^T W_i + \sum_{j \in N_i} e_{ij} \times \\
& (\mathrm{d}_j + \mathrm{e}_{j0}) (\psi_j^* \varsigma_j + \varepsilon_{g_j}) R_{ii}^{-1} (\psi_j^* \varsigma_j + \varepsilon_{g_j})^T \times \\
& \nabla \sigma_j^T W_j] - \frac{1}{2} (\mathrm{d}_i + \mathrm{e}_{i0})^2 W_i^T \nabla \sigma_i (\psi_i^* \varsigma_i + \varepsilon_{g_i}) \times \\
& R_{ii}^{-T} (\psi_i^* \varsigma_i + \varepsilon_{g_i})^T \nabla \sigma_i^T W_i - \frac{1}{2} \sum_{j \in N_i} (\mathrm{d}_j + \mathrm{e}_{j0})^2 \times \\
& W_j^T \nabla \sigma_j (\psi_j^* \varsigma_j + \varepsilon_{g_j}) R_{jj}^{-T} R_{ij} R_{jj}^{-1} \times \\
& (\psi_j^* \varsigma_j + \varepsilon_{g_j})^T \nabla \sigma_j^T W_j + e_{B_i},
\end{aligned}$$

ثابت مثبت b_δ به اندازه کافی بزرگ وجود دارد به قسمی که

$$\sum_{i=1}^{N} \|\omega_{i0}\| \le \sum_{i=1}^{N} \{b_{\nabla\omega_i} b_f b_\delta \|\delta_i\| + b_{\nabla\omega_i} \times (d_i + e_{i0})^2 \|g_i(x_i)\| \sigma_{\min}(R_{ii}) (\|\psi_i^* \varsigma_i\| + \|\tilde{\psi}_i \varsigma_i\|) b_{\nabla\sigma_i} (\|W_i\| + \|\tilde{W}_{i+N}\|) + + b_{\nabla\omega_i} \times \sum_{j\in N_i} e_{ij}(d_j + e_{j0}) \|g_j(x_j)\| \sigma_{\min}(R_{jj}) \times b_{\nabla\sigma_j} (\|\tilde{\psi}_j \varsigma_j\| + \|\psi_j^* \varsigma_j\|)(\|W_j\| + \|\tilde{W}_{j+N}\|)\}$$

برای دومین جمله $\dot{L}$ ، $\sum_{i=1}^{N} \dot{L}_i = -\sum_{i=1}^{N} \tilde{W}_i^T \alpha_i^{-1} \dot{\hat{W}}_i$ ،

با استفاده از (۲۸)، (۳۴)-(۳۸)، (۱۸) و $g_i = \psi_i^* \varsigma_i + \varepsilon_{g_i}$ داریم

$$\dot{L}_i = -\tilde{W}_i^T \alpha_i^{-1} \dot{\hat{W}}_i = \tilde{W}_i^T \frac{B_{i+N}}{(1 + B_{i+N}{}^T B_{i+N})^2} \times \{W_i^T \nabla\sigma_i(\tilde{\varphi}_i z_i(x_i, x_j, u_{i+N}, u_{-i+N})) - B_{i+N}{}^T \tilde{W}_i + \frac{1}{2}(W_i - \tilde{W}_{i+N})^T \hat{D}_i (W_i - \tilde{W}_{i+N}) + \frac{1}{2} \sum_{j\in N_i} (W_j - \tilde{W}_{j+N})^T \hat{E}_j (W_j - \tilde{W}_{j+N}) - W_i^T \frac{\partial\sigma_i}{\partial\delta_i}((d_i + e_{i0})(\psi_i^* \varsigma_i)(u_i - \hat{u}_i) - \sum_{j\in N_i} e_{ij}(\psi_j^* \varsigma_j)(u_j - \hat{u}_j)) - W_i^T \nabla\sigma_i \varepsilon_{T_i} + \varepsilon_{L_i} + e_{B_i}, \quad (۴۳)$$

$$\varepsilon_{L_i} = -\frac{1}{2}(\mathrm{d}_i + \mathrm{e}_{i0})^2 W_i^T \nabla\sigma_i(\psi_i^* \varsigma_i + \varepsilon_{g_i}) \times R_{ii}^{-T}(\psi_i^* \varsigma_i + \varepsilon_{g_i})^T \nabla\sigma_i^T W_i - \frac{1}{2} \sum_{j\in N_i} (\mathrm{d}_j + \mathrm{e}_{j0})^2 W_j^T \nabla\sigma_j(\psi_j^* \varsigma_j + \varepsilon_{g_j}) \times R_{jj}^{-T} R_{ij} R_{jj}^{-1}(\psi_j^* \varsigma_j + \varepsilon_{g_j})^T \nabla\sigma_j^T W_j, \quad (۴۴)$$

که در آن برای $i = 1, \ldots, \mathrm{N}$ داریم

$$c_j = (\mathrm{d}_j + \mathrm{e}_{j0}), \quad c_i = (\mathrm{d}_i + \mathrm{e}_{i0}),$$
$$D_i = c_i^2 \nabla\sigma_i(\psi_i^* \varsigma_i) R_{ii}^{-T} (\psi_i^* \varsigma_i)^T \nabla\sigma_i^T,$$
$$\tilde{D}_i = c_i^2 \nabla\sigma_i(\tilde{\psi}_i \varsigma_i) R_{ii}^{-T} (\tilde{\psi}_i \varsigma_i)^T \nabla\sigma_i^T,$$
$$E_j = c_j^2 \nabla\sigma_j(\psi_j^* \varsigma_j) R_{jj}^{-T} R_{ij} R_{jj}^{-1} (\psi_j^* \varsigma_j)^T \nabla\sigma_j^T,$$
$$\tilde{E}_j = c_j^2 \nabla\sigma_j(\tilde{\psi}_j \varsigma_j) R_{jj}^{-T} R_{ij} R_{jj}^{-1} (\tilde{\psi}_j \varsigma_j)^T \nabla\sigma_j^T,$$
$$\hat{D}_i = \tilde{D}_i + D_i + c_i^2 \nabla\sigma_i(\tilde{\psi}_i \varsigma_i) R_{ii}^{-T} (\psi_i^* \varsigma_i)^T \nabla\sigma_i^T + c_i^2 \nabla\sigma_i(\psi_i^* \varsigma_i) R_{ii}^{-T} (\tilde{\psi}_i \varsigma_i)^T \nabla\sigma_i^T,$$
$$\hat{E}_j = \tilde{E}_j + E_j + c_j^2 \nabla\sigma_j \tilde{\psi}_j \varsigma_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} \varsigma_j^T \psi_j^{*T} \times \nabla\sigma_j^T + c_j^2 \nabla\sigma_j \psi_j^* \varsigma_j R_{jj}^{-T} R_{ij} R_{jj}^{-1} \varsigma_j^T \tilde{\psi}_j^T \nabla\sigma_j^T.$$

با توجه به اینکه $\hat{D}_i, \hat{E}_j, i = 1, \ldots, \mathrm{N}, \mathrm{j} \in \mathrm{N}_i$ ، کراندار هستند. بنابراین جملات

$$\frac{1}{4} W_i^T D_i W_i, \frac{1}{4} W_i^T \hat{D}_i W_i, \frac{1}{4} W_j^T E_j W_j, \frac{1}{4} W_j^T \hat{E}_j W_j$$

که در $i = 1, \ldots, \mathrm{N}, \dot{L}_i(\mathrm{t})$ ظاهر می شوند نیز به صورت زیر کراندار هستند و در (۴۵) $k_{H_i}, k_{G_i}, k_{S_j}, k_{\Upsilon_j}, k_{N_i}$ و k_{O_j} ثابت های اسکالر هستند.

$$\left\|\frac{1}{2} W_i^T D_i W_i\right\| \le k_{H_i}, \quad \left\|\frac{1}{2} W_i^T \hat{D}_i W_i\right\| \le k_{G_i},$$
$$\left\|\frac{1}{2} W_j^T E_j W_j\right\| \le k_{S_j}, \left\|\frac{1}{2} W_j^T \hat{E}_j W_j\right\| \le k_{\Upsilon_j}, \quad (۴۵)$$
$$\left\|\hat{D}_i\right\| \le k_{N_i}, \quad \left\|\hat{E}_j\right\| \le k_{\mathrm{O}_j}.$$

با استفاده از (۴۵)، (۳۵)-(۳۷) و نامساوی یونگ (لم ۳) و فرض های ۱ و ۲، $\dot{L}_i(t)$ به صورت زیر در می آید

$$\dot{L}_i(t) \le \dot{L}\_i + \tilde{W}_{i+N}{}^T \hat{D}_i \tilde{W}_{i+N} \frac{\bar{B}_{i+N}{}^T}{2m_{s_{i+N}}} \tilde{W}_i + \sum_{j\in N_i} \tilde{W}_{j+N}{}^T \hat{E}_j \tilde{W}_{j+N} \frac{\bar{B}_{i+N}{}^T}{2m_{s_{i+N}}} \tilde{W}_i \quad (۴۶)$$

که در آن

$$\dot{L}\_i = -r_i \tilde{W}_i^T \bar{B}_{i+N} \bar{B}_{i+N}^T \tilde{W}_i + \left\| \tilde{W}_i^T \frac{\bar{B}_{i+N}}{m_{s_{i+N}}} \right\| \times$$
$$(a_{i_1} + k_{T_i}) + \frac{\varepsilon}{2} a_{i_2} \delta_i^T \delta_i - \frac{k_{N_i}}{2} W_i \tilde{W}_{i+N} \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i$$
$$- \frac{k_{N_i}}{4} \tilde{W}_{i+N} W_i \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i - \sum_{j \in N_i} \frac{k_{O_j}}{4} W_j^T \tilde{W}_{j+N} \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i$$
$$- \sum_{j \in N_i} \frac{k_{O_j}}{4} \tilde{W}_{j+N}^T W_j \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i + (\mathrm{d}_i + \mathrm{e}_{i0})^2 W_i^T \nabla \sigma_i \times$$
$$(\tilde{\psi}_i \varsigma_i) R_{ii}^{-1} (\psi_i^* \varsigma_i)^T \nabla \sigma_i^T (\tilde{W}_{i+N} - W_i) \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i$$
$$- W_i^T \tilde{D}_i W_i \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i + W_i^T \tilde{D}_i \tilde{W}_{i+N} \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i$$
$$+ W_i^T \nabla \sigma_i \sum_{j \in N_i} (\mathrm{d}_j + \mathrm{e}_{j0}) \mathrm{e}_{ij} (\tilde{\psi}_j \varsigma_j) R_{jj}^{-1} (\psi_j^* \varsigma_j)^T \nabla \sigma_j^T \times$$
$$(W_j - \tilde{W}_{j+N}) \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i + W_i^T \nabla \sigma_i \sum_{j \in N_i} (\mathrm{d}_j + \mathrm{e}_{j0}) \mathrm{e}_{ij} \times$$
$$(\tilde{\psi}_j \varsigma_j) R_{jj}^{-1} (\tilde{\psi}_j \varsigma_j)^T \nabla \sigma_j^T (W_j - \tilde{W}_{j+N}) \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i$$
$$+ (e_{B_i} + \varepsilon_{L_i} - W_i^T \nabla \sigma_i \varepsilon_{T_i}) \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i - W_i^T \nabla \sigma_i [\psi_i^* \varsigma_i \times$$
$$(\mathrm{d}_i + \mathrm{e}_{i0})(-(\mathrm{d}_i + \mathrm{e}_{i0}) R_{ii}^{-1} (\psi_i^* \varsigma_i)^T \nabla \sigma_i^T W_i - (\mathrm{d}_i + \mathrm{e}_{i0}) \times$$
$$R_{ii}^{-1} (\varepsilon_{g_i})^T \nabla \sigma_i^T W_i + (\mathrm{d}_i + \mathrm{e}_{i0}) R_{ii}^{-1} ((\tilde{\psi}_i + \psi_i^*) \varsigma_i)^T \times$$
$$\nabla \sigma_i^T W_i - (\mathrm{d}_i + \mathrm{e}_{i0}) R_{ii}^{-1} ((\tilde{\psi}_i + \psi_i^*) \varsigma_i)^T \nabla \sigma_i^T \tilde{W}_{i+N})$$
$$- \sum_{j \in N_i} \mathrm{e}_{ij} (\psi_j^* \varsigma_j + \varepsilon_{g_j}) (-(\mathrm{d}_j + \mathrm{e}_{j0}) R_{jj}^{-1} (\psi_j^* \varsigma_j)^T \nabla \sigma_j^T W_j$$
$$- (\mathrm{d}_j + \mathrm{e}_{j0}) R_{jj}^{-1} (\varepsilon_{g_j})^T \nabla \sigma_j^T W_j + (\mathrm{d}_j + \mathrm{e}_{j0}) R_{jj}^{-1} \times$$
$$((\tilde{\psi}_j + \psi_j^*) \varsigma_j)^T \nabla \sigma_j^T W_j - (\mathrm{d}_j + \mathrm{e}_{j0}) R_{jj}^{-1} \times$$
$$((\tilde{\psi}_j + \psi_j^*) \varsigma_j)^T \nabla \sigma_j^T \tilde{W}_{j+N})] \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i,$$

$$\left\| W_i^T \nabla \sigma_i \left(\sum_{j \in N_i} e_{ij} (\tilde{\theta}_i \xi_i - \tilde{\theta}_j \xi_j) + e_{i0} (\tilde{\theta}_i \xi_i - \tilde{\theta}_0 \xi_0) \right) \right\| \le$$
$$a_{i_1} + a_{i_2} \|\delta_i\|, \; k_{T_i} = k_{G_i} + \sum_{j \in N_i} k_{\Upsilon_j}, \; r_i = 1 - \frac{a_{i_2}}{\varepsilon m_{s_{i+N}}}.$$

با استفاده از جملات (۴۲)، (۴۶) ، (۱۸) و (۳۵)- (۳۷)، (۴۱) به صورت زیر در می آید

$$\dot{L}(t) \le \dot{L}\_V_i + \sum_{i=1}^{N} \Big\{ \dot{L}\_i + \tilde{W}_{i+N}^T \hat{D}_i \times$$
$$\tilde{W}_{i+N} \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} W_i - \tilde{W}_{i+N}^T \hat{D}_i W_i \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} W_i$$
$$+ \tilde{W}_{i+N}^T \hat{D}_i W_i \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} \tilde{W}_i - \frac{k_{N_i}}{2} W_i \tilde{W}_{i+N} \times$$
$$\frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} \tilde{W}_i + \sum_{j \in N_i} \tilde{W}_{j+N}^T \hat{E}_j \tilde{W}_{j+N} \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} W_i$$
$$- \sum_{j \in N_i} \tilde{W}_{j+N}^T \hat{E}_j W_j \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} W_i + \sum_{j \in N_i} \tilde{W}_{j+N}^T \times$$
$$\hat{E}_j W_j \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} \tilde{W}_i \Big\} + \sum_{i=1}^{N} L_{W_{i+N}} \;, \quad (۴۷)$$

$$\sum_{i=1}^{N} L_{W_{i+N}} = - \sum_{i=1}^{N} \tilde{W}_{i+N}^T \alpha_{i+N}^{-1} \dot{\hat{W}}_{i+N}$$
$$+ \sum_{i=1}^{N} \tilde{W}_{i+N}^T \hat{D}_i \hat{W}_{i+N} \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} \hat{W}_i$$
$$+ \underbrace{\sum_{i=1}^{N} \sum_{j \in i_N} \tilde{W}_{j+N}^T \hat{E}_j \hat{W}_{j+N} \frac{\bar{B}_{j+N}^T}{2m_{s_j}} \hat{W}_j}_{\sum_{i=1}^{N} \tilde{W}_{i+N}^T \sum_{j \in N_i} \hat{E}_j \hat{W}_{i+N} \frac{\bar{B}_{j+N}^T}{2m_{s_j}} \hat{W}_j} \Big)$$

اکنون با استفاده از $L_{W_{i+N}}$ ، $\dot{\hat{W}}_{i+N}$ به صورت زیر به دست می آید:

$$\dot{\hat{W}}_{i+N} = -\alpha_{i+N} \{ (S_i \hat{W}_{i+N} - F_i \hat{W}_i) - \hat{D}_i \hat{W}_{i+N} \times$$
$$\frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} \hat{W}_i - \sum_{j \in i_{N_i}} \hat{E}_j \hat{W}_{i+N} \frac{\bar{B}_{j+N}^T}{2m_{s_j}} \hat{W}_j \} \quad (۴۸)$$

با به کارگیری (۴۸) در (۴۷)، $L_{W_{i+N}}$ با جملات زیر جایگزین می شود.

$$\tilde{W}_{i+N}^T S_i W_i - \tilde{W}_{i+N}^T S_i \tilde{W}_{i+N}$$
$$- \tilde{W}_{i+N}^T F_i W_i + \tilde{W}_{i+N}^T F_i \tilde{W}_i \quad (۴۹)$$

از آنجایی که $Q_i(\delta_i) > 0, \; i = 1, \ldots, N$ ، وجود دارد $\forall i, q_i > 0$ به طوری که $\delta_i^T q_i \delta_i < Q_i(\delta_i)$ ، بنابراین $-\delta_i^T q_i \delta_i > -Q_i(\delta_i)$.

$\dot{L}$ به صورت زیر نوشته می شود

$$m_{44}^i = -\sum_{j\in N_i} \hat{E}_j \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} W_i,\quad d_1^i = b_{\nabla\omega_i} b_f b_\delta ,$$

$$d_2^i = \frac{\bar{B}_{i+N}}{m_{s_i}}(\mathrm{a}_{i_1} + \mathrm{k}_{T_i}) - (\mathrm{d}_i + \mathrm{e}_{i0})^2 W_i^T \nabla\sigma_i(\tilde{\psi}_i\varsigma_i)\times$$
$$R_{ii}^{-1}(\psi_i^*\varsigma_i)^T \nabla\sigma_i^T W_i \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} - W_i^T \tilde{D}_i W_i \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}}$$
$$-W_i^T \nabla\sigma_i \sum_{j\in N_i} (\mathrm{d}_j + \mathrm{e}_{j0})\mathrm{e}_{ij}(\tilde{\psi}_j\varsigma_j) R_{jj}^{-1}(\psi_j^*\varsigma_j)^T \times$$
$$\nabla\sigma_j^T W_j \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} - W_i^T \nabla\sigma_i \sum_{j\in N_i} (\mathrm{d}_j + \mathrm{e}_{j0})\mathrm{e}_{ij}(\tilde{\psi}_j\varsigma_j)\times$$
$$R_{jj}^{-1}(\tilde{\psi}_j\varsigma_j)^T \nabla\sigma_j^T W_j \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} - W_i^T \nabla\sigma_i[\psi_i^*\varsigma_i \times$$
$$(-(\mathrm{d}_i + \mathrm{e}_{i0}) R_{ii}^{-1}(\psi_i^*\varsigma_i + \varepsilon_{g_i})^T \nabla\sigma_i^T W_i + (\mathrm{d}_i + \mathrm{e}_{i0})\times$$
$$R_{ii}^{-1}((\tilde{\psi}_i + \psi_i^*)\varsigma_i)^T \nabla\sigma_i^T W_i) - \sum_{j\in N_i} \mathrm{e}_{ij}(\psi_j^*\varsigma_j + \varepsilon_{g_j})\times$$
$$(-(\mathrm{d}_j + \mathrm{e}_{j0}) R_{jj}^{-1}(\psi_j^*\varsigma_j + \varepsilon_{g_j})^T \nabla\sigma_j^T W_j$$
$$+(\mathrm{d}_j + \mathrm{e}_{j0}) R_{jj}^{-1}((\tilde{\psi}_j + \psi_j^*)\varsigma_j)^T \nabla\sigma_j^T W_j)] \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}}$$
$$+(e_{B_i} + \varepsilon_{L_i} - W_i^T \nabla\sigma_i \varepsilon_{T_i}) \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}},$$
$$d_3^i = -W_i^T \nabla\sigma_i (\mathrm{d}_i + \mathrm{e}_{i0})^2 (\psi_i^*\varsigma_i + \varepsilon_{g_i}) R_{ii}^{-1} \times$$
$$((\tilde{\psi}_i + \psi_i^*)\varsigma_i)^T \nabla\sigma_i^T - \hat{D}_i W_i \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} W_i$$
$$+W_i^T S_i - W_i^T F_i + b_{\nabla\omega_i}(d_i + e_{i0})^2 \left\| g_i(x_i) \right\| \times$$
$$\sigma_{\min}(R_{ii})(\left\| \psi_i^*\varsigma_i \right\| + \left\| \tilde{\psi}_i\varsigma_i \right\|) b_{\nabla\sigma_i} \left\| \tilde{W}_{i+N} \right\|,$$
$$d_4^i = W_i^T \nabla\sigma_i (\sum_{j\in N_i} e_{ij}((\mathrm{d}_j + \mathrm{e}_{j0})(\psi_j^*\varsigma_j + \varepsilon_{g_j}) \times$$
$$R_{jj}^{-1}((\tilde{\psi}_j + \psi_j^*)\varsigma_j)^T \nabla\sigma_j^T)) - \sum_{j\in N_i} \hat{E}_j W_j \times$$
$$\frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} W_i + b_{\nabla\omega_i} \sum_{j\in N_i} e_{ij}(d_j + e_{j0}) \left\| g_j(x_j) \right\| \times$$
$$\sigma_{\min}(R_{jj}) b_{\nabla\sigma_j} (\left\| \tilde{\psi}_j\varsigma_j \right\| + \left\| \psi_j^*\varsigma_j \right\|) \left\| \tilde{W}_{j+N} \right\|,$$

$$\dot{L}(t) = \sum_{i=1}^{N} \{C_i - \tilde{Z}_i^T M_i \tilde{Z}_i + D_i \tilde{Z}_i\}, \qquad (۵۰)$$
$$\tilde{Z}_i = \left[\delta_i\ , \tilde{W}_i\ , \tilde{W}_{i+N}\ , \tilde{W}_{j+N}\right]^T$$
$$, D_i = \left[d_1^i,\ d_2^i,\ d_3^i,\ d_4^i\right] \text{ که در آن}$$
$$, M_i = \begin{bmatrix} m_{11}^i & m_{12}^i & m_{13}^i & m_{14}^i \\ m_{21}^i & m_{22}^i & m_{23}^i & m_{24}^i \\ m_{31}^i & m_{32}^i & m_{33}^i & m_{34}^i \\ m_{41}^i & m_{42}^i & m_{43}^i & m_{44}^i \end{bmatrix}$$
$$m_{12}^i = m_{13}^i = m_{14}^i = m_{41}^i = m_{31}^i$$
$$= m_{21}^i = 0,\ m_{34}^i = m_{43}^i = 0,$$

$$m_{11}^i = \frac{1}{2} q_i - \frac{\varepsilon}{2} a_{i_4} I\ ,\quad m_{22}^i = r_i \bar{B}_{i+N} \bar{B}_{i+N}^T ,$$

$$m_{23}^i = -\frac{1}{4m_{s_{i+N}}} \hat{D}_i W_i \bar{B}_{i+N}^T + \frac{k_{N_i}}{2} W_i \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}}$$
$$-\frac{1}{2}(\mathrm{d}_i + \mathrm{e}_{i0})^2 W_i^T \nabla\sigma_i(\tilde{\psi}_i\varsigma_i) R_{ii}^{-1}(\psi_i^*\varsigma_i)^T \times$$
$$\nabla\sigma_i^T \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} - \frac{1}{2} W_i^T \tilde{D}_i \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}}$$
$$-\frac{1}{2} W_i^T \nabla\sigma_i(\psi_i^*\varsigma_i)(\mathrm{d}_i + e_{i0}) R_{ii}^{-1} \times$$
$$((\tilde{\psi}_i + \psi_i^*)\varsigma_i)^T \nabla\sigma_i^T \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} - \frac{1}{2} F_i = m_{32}^{i\,T} ,$$

$$m_{24}^i = -\frac{1}{4} \sum_{j\in N_i} \hat{E}_j W_j \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} + \frac{1}{2} \sum_{j\in N_i} k_{O_j} \times$$
$$W_j \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} - \frac{1}{2} W_i^T \nabla\sigma_i \sum_{j\in N_i} (\mathrm{d}_j + \mathrm{e}_{j0})\mathrm{e}_{ij} \times$$
$$(\tilde{\psi}_j\varsigma_j) R_{jj}^{-1}(\psi_j^*\varsigma_j)^T \nabla\sigma_j^T \frac{\bar{B}_{i+N}^T}{m_{s_i}} - \frac{1}{2} W_i^T \times$$
$$\nabla\sigma_i \sum_{j\in N_i} (\mathrm{d}_j + \mathrm{e}_{j0})\mathrm{e}_{ij}(\tilde{\psi}_j\varsigma_j) R_{jj}^{-1}(\tilde{\psi}_j\varsigma_j)^T \times$$
$$\nabla\sigma_j^T \frac{\bar{B}_{i+N}^T}{m_{s_i}} - \frac{1}{2} W_i^T \nabla\sigma_i \sum_{j\in N_i} \mathrm{e}_{ij}(\psi_j^*\varsigma_j) \times$$
$$(d_j + e_{j0}) R_{jj}^{-1}((\tilde{\psi}_j + \psi_j^*)\varsigma_j)^T \nabla\sigma_j^T \frac{\bar{B}_{i+N}^T}{m_{s_{i+N}}} = m_{24}^{i\,T},$$

$$m_{33}^i = -\hat{D}_i \frac{\bar{B}_{i+N}^T}{2m_{s_{i+N}}} W_i + S_i\ ,$$

$$C_i = \dot{\bar{L}}_{V_i} - W_i^T \nabla\sigma_i\{(\mathrm{d}_i+\mathrm{e}_{i0})^2(\psi_i^*\varsigma_i)R_{ii}^{-1}\times$$
$$((\tilde{\psi}_i+\psi_i^*)\varsigma_i)^T\nabla\sigma_i^T W_i + (\mathrm{d}_i+\mathrm{e}_{i0})^2(\varepsilon_{g_i})\times$$
$$R_{ii}^{-1}((\tilde{\psi}_i+\psi_i^*)\varsigma_i)^T\nabla\sigma_i^T W_i) - \sum_{j\in N_i} e_{ij}\times$$
$$((\mathrm{d}_j+\mathrm{e}_{j0})(\psi_j^*\varsigma_j)R_{jj}^{-1}((\tilde{\psi}_j+\psi_j^*)\varsigma_j)^T\nabla\sigma_j^T W_j$$
$$+(\mathrm{d}_j+\mathrm{e}_{j0})(\varepsilon_{g_j})R_{jj}^{-1}((\tilde{\psi}_j+\psi_j^*)\varsigma_j)^T\nabla\sigma_j^T W_j)\}$$
$$+b_{\nabla\omega_i}(d_i+e_{i0})^2\left\|g_i(x_i)\right\|\sigma_{\min}(R_{ii})(\left\|\psi_i^*\varsigma_i\right\|$$
$$+\left\|\tilde{\psi}_i\varsigma_i\right\|)b_{\nabla\sigma_i}\left\|W_i\right\|++b_{\nabla\omega_i}\sum_{j\in N_i}e_{ij}(d_j+e_{j0})\times$$
$$\left\|g_j(x_j)\right\|\sigma_{\min}(R_{jj})b_{\nabla\sigma_j}(\left\|\tilde{\psi}_j\varsigma_j\right\|+\left\|\psi_j^*\varsigma_j\right\|)\left\|W_j\right\|,$$

و $C_i \le C_{i\max}$ و $D_i \le D_{i\max}$ که ثابت های $C_{i\max}$ و $D_{i\max}$ اسکالر می باشند. پارامترهای تنظیم باید به گونه ای انتخاب شوند به طوری که ماتریس M_i مثبت معین باشد.

$$M_i = \begin{bmatrix} \frac{1}{2}q_i & 0 & 0 \\ 0 & I & M_{23}^i \\ 0 & M_{32}^i & M_{33}^i \end{bmatrix}, M_{23}^i = \begin{bmatrix} m_{23}^i & m_{24}^i \end{bmatrix}, \quad (۵۱)$$
$$M_{32}^i = \begin{bmatrix} m_{32}^i \\ m_{42}^i \end{bmatrix}, M_{33}^i = \begin{bmatrix} m_{33}^i & 0 \\ 0 & m_{44}^i \end{bmatrix}$$

برای مثبت معین بودن M_i باید ویژگی های زیر برقرار باشند:

الف) $q_i > 0$. ب) $I > 0$. ج) مکمل شور[۱] برای I ، $D_{22}^i = I - M_{23}^i M_{33}^{i\,-1} M_{32}^i > 0$ می باشد که با انتخاب مناسب S_i و F_i می تواند برقرار باشد. پ) مکمل شور برای M_{33}^i ، $D_{33}^i = M_{33}^i - M_{32}^i I^{-1} M_{23}^i > 0$ می باشد که با انتخاب مناسب S_i و F_i می تواند برقرار باشد.

اکنون برای (۵۰) داریم

$$\dot{L} < \sum_{i=1}^{N}\{-\left\|\tilde{Z}_i\right\|^2\sigma_{\min}(M_i) + D_{i\max}\left\|\tilde{Z}_i\right\| + C_{i\max}\} \quad (۵۲)$$

که $\sigma_{\min}(M_i)$ کوچکترین مقدار ویژه M_i را نشان می دهد. با کامل کردن مربعات، مشتق لیاپانوف منفی خواهد بود اگر

$$\left\|\tilde{Z}_i\right\| > \frac{D_{i\max}}{2\sigma_{\min}(M_i)} + \sqrt{\frac{D_{i\max}^{\ 2}}{4\sigma_{\min}^{\ 2}(M_i)}\left\|\tilde{Z}_i\right\| + \frac{C_{i\max}}{\sigma_{\min}(M_i)}} \equiv B_{\tilde{Z}_i} \quad (۵۳)$$

مشاهده می شود که اگر (۵۳) از یک کران مشخص فراتر رود، آنگاه $\dot{L}$ منفی خواهد بود. بنابراین مطابق با توسعه لیاپانوف استاندارد، تحلیل های بالا نشان می دهد که حالت خطاهای محلی و خطای وزن ها کراندار نهایی یکنواخت هستند [۳۶]. شرط (۵۳) برقرار است اگر نرم هر قسمت از $\tilde{Z}_i$ از کران $B_{\tilde{Z}_i}$ فراتر رود، یعنی $\delta_i > B_{\tilde{Z}_i}$ ، $\tilde{W}_{i+N} > B_{\tilde{Z}_i}$ ، $\tilde{W}_{j+N} > B_{\tilde{Z}_i}$. به این ترتیب اثبات کامل می گردد. ■

قضیه ۳ (همگرایی به تعادل نش)- در نظر بگیرید که تمام شرایط و فرض های قضیه ۲ برقرار باشد، آنگاه الف) شبکه عصبی نقاد و عملگر به پاسخ معادله CHJ تقریبی همگرا می‌شوند یا به عبارت دیگر وزن های تخمین $\hat{W}_i$ به وزن های بهینه W_i همگرا می‌شوند. ب) مجموعه ورودی‌های $\hat{u}_i, i = 1,...,N$ به حل تعادل نش تقریبی بازی گرافی دیفرانسیلی برای سیستم های خطی با دینامیک نامعلوم همگرا می‌شوند.

اثبات قضیه ۳- اثبات یک نتیجه مستقیم از قضیه ۲ می باشد. ■

۵- نتایج شبیه‌سازی

در این بخش برای ارزیابی روش معرفی شده، الگوریتم پیشنهادی به یک مثال شبیه‌سازی اعمال می گردد. یک سیستم چندعاملی دارای ۵ گره، مطابق با شکل ۱ را در نظر بگیرید. وزن های اتصال و وزن یال ها، ۱ در نظر گرفته شده اند. دینامیک عامل ها به صورت زیر می باشد.

$$\dot{x}_0 = \begin{pmatrix} p_1 x_{02} \\ p_2 x_{01} + p_3 x_{02} + p_4 x_{02} x_{01}^{\ 2} \end{pmatrix},$$
$$\dot{x}_i = f_i(x_i) + g_i(x_i)\,u_i =$$
$$\begin{pmatrix} p_1 x_{i2} \\ p_2 x_{i1} + p_3 x_{i2} + p_4 x_{i2} x_{i1}^{\ 2} \end{pmatrix} + \begin{bmatrix} 0 \\ p_{i+4} x_{i1} x_{i2} \end{bmatrix} u_i,$$
$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \end{bmatrix}, \ i = 1,2,...,5.$$

بردار پارامترهای سیستم چندعاملی $P = [p_1, p_2, ..., p_9]$ نامعلوم است، به طوری که مقادیر واقعی پارامترها برابر با $P = [1, -1, 0.5, -0.5, -0.8, 1, 0.5, -0.2, 1.4]$ در نظر گرفته شده اند. در الگوریتم شناساگرهای شبکه های عصبی، $a = 40,\ p = 1000,\ \Gamma = 70$ و ثابت نمونه برداری برابر

[۱] Schur Complement

با 0.001 ثانیه می باشد. برای $i, j = 1,...,5$ ، در نظر بگیرید

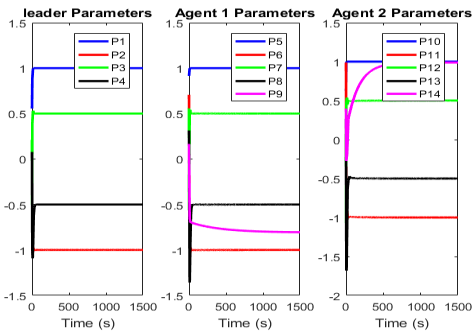
$$Q_i(\delta_i) = \delta_i^T Q_{ii} \delta_i = \delta_i \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \delta_i ,$$

$R_{ii} = 10,\ R_{ij} = 1,\ (\mathrm{i} \neq \mathrm{j},\ \mathrm{j} \in \mathrm{N}_i)$. در الگوریتم به کار رفته، پارامترهای طراحی به صورت $S_i = F_i = 100I$ و $\alpha_i = 1$ ، $i = 1,...,5$ انتخاب شده اند. بردار اطلاعات در دسترس برای هر عامل، $\delta_i = [\delta_{i1}, \delta_{i2}]^T$, $i = 1,...,5$ می باشد که توسط گراف ارتباطی محدود شده است. توابع فعالیت شبکه های عصبی دینامیک خطای عامل ها به صورت

$$\sigma_i = [\delta_{i1}{}^2, \delta_{i1}\delta_{i2}, \delta_{i2}{}^2, \delta_{i1}{}^3, \delta_{i1}{}^2\delta_{i2}, \delta_{i1}\delta_{i2}{}^2, \delta_{i2}{}^3, \delta_{i1}{}^4, \delta_{i1}{}^3\delta_{i2}, \delta_{i1}{}^2\delta_{i2}{}^2, \delta_{i1}\delta_{i2}{}^3, \delta_{i2}{}^4],$$

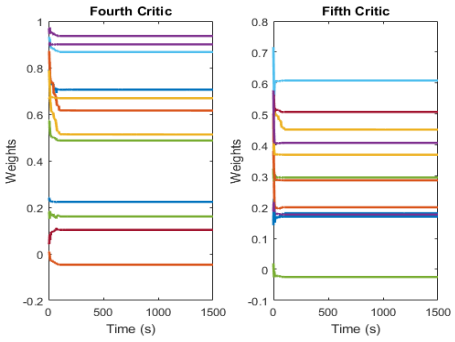
انتخاب می شوند. هر $\sigma_i, i = 1,...,5$ شامل توان های δ_{i1} و δ_{i2} از درجه ۴ می باشد. توجه شود که یک نویز مجموع سینوسی کوچک از ابتدا تا ثانیه ۱۴۰۰ به ورودی های کنترلی اضافه می شود تا شرط تحریک پایا بودن تضمین شود. الگوریتم یادگیری توزیع شده بهینه معرفی شده در قضیه ۲ برای حل مسئله اجماع، به بازی گرافی دیفرانسیلی غیرخطی تحت دینامیک نامعلوم اعمال شده است. شکل ۲ نشان می دهد که چگونه پارامترهای نامعلوم سیستم چندعاملی (پارامترهای رهبر، عامل های ۱ و۲ به عنوان نمونه آورده شده اند.) به مقادیر واقعی خود همگرا می شود، به گونه ای که نهایتا پارامترهای سیستم چندعاملی به مقادیر زیر همگرا می گردد.

$$P = [1, -1, 0.5, -0.5, -0.79, 0.98, 0.45, -0.21, 1.37]$$



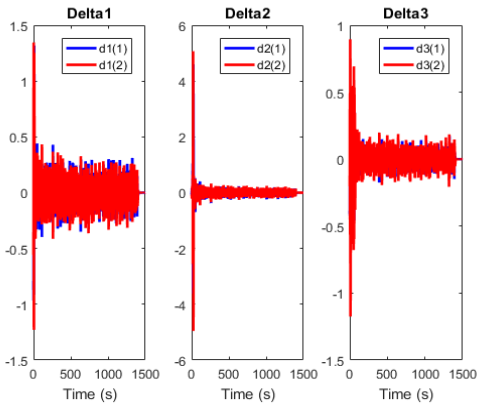
شکل ۲. همگرایی پارامترهای عامل های غیرخطی ۱ و ۲ و عامل رهبر

شکل ۳ نیز همگرایی وزن های شبکه های نقاد را نمایش می دهد (وزن های شبکه های نقاد، عامل های ۴ و۵ به عنوان نمونه آورده شده اند.).



شکل ۳. همگرایی وزن های شبکه های نقاد عامل های غیرخطی ۴ و ۵

شکل ۴ خطای ردیابی محلی برای عامل ها را نشان می دهند که پس از حذف نویز ورودی ها در ثانیه ۱۴۰۰، خطای ردیابی عامل ها تقریبا به صفر همگرا شده اند (خطای ردیابی ، عامل های ۱، ۲ و ۳ به عنوان نمونه آورده شده اند.). با توجه به همگرایی خطا به صفر در شکل های مذکور، همزمان سازی همه عامل ها به رهبر نشان داده شده است.



شکل ۴. خطای ردیابی عامل های غیرخطی ۱، ۲ و ۳

نتایج نشان می دهد که الگوریتم کنترل بهینه توزیع شده پیشنهادی برای همزمانسازی سیستم های چندعاملی غیر خطی زمان پیوسته با دینامیک نامعلوم، به حل تقریبی بهینه همگرا شده است.

۸- نتیجه گیری

در این مقاله الگوریتم کنترل بهینه توزیع شده برخطی بر اساس تکنیک برنامه ریزی پویای تقریبی برای حل بهینه برخط مسئله همزمانسازی سیستم های چندعاملی غیرخطی دارای دینامیک نامعلوم ارایه شده است. حین تقریب سیاست های بهینه، دینامیک نامعلوم عامل ها نیز به طور همزمان شناسایی گردید. الگوریتم یادگیری توزیع شده، با استفاده از ساختار شبکه های عصبی عملگر-نقاد جهت تقریب سیاست های بهینه بازیکنان ارایه شد. برای شناسایی دینامیک نامعلوم عامل ها، شناساگرهای شبکه های عصبی به صورت هم‌زمان در کنار شبکه های عملگر-نقاد برای هر عامل به کار گرفته شده اند. کرانداری سیگنال های حلقه بسته توسط پایداری لیاپانوف اثبات شده و نشان داده شد که سیاست های به دست

آمده تعادل نش تقریبی بازی را نتیجه می دهند. صحت الگوریتم پیشنهاد شده نیز در شبیه سازی مورد بررسی قرار گرفت.

مراجع

[1] Hong Y., Hu J., Gao L., 2006 "Tracking control for multi-agent consensus with an active leader and variable topology," Automatica, 42 (7), 1177–1182.
[2] Ren W., Moore K., Chen Y., 2007, "High-order and model reference consensus algorithms in cooperative control of multivehicle systems," J. Dynam. Syst., Meas., Control, 129(5), 678–688.
[3] Wang X., Chen G., 2002, "Pinning control of scale-free dynamical networks," Physica A, 310(3-4), 521–531.
[4] Wu Y., Meng X., Xie L., Lu R., Su H., Wu Z. G., 2017, "An input-based triggering approach to leader-following problems," Automatica, 75, 221–228.
[5] Zhang D., Xu Z., Wang Q. G., Zhao Y. B., 2017, "Leader–follower H_∞ consensus of linear multi-agent systems with aperiodic sampling and switching connected topologies," ISA Transactions, 68, 150–159.
[6] Wang B., Wang J., Zhang B., Lin H., Li X., Wang H., 2016, "Leader–follower consensus for multi-agent systems with three-layer network framework and dynamic interaction jointly connected topology," Neurocomputing, 207 (26), 231-239.
[7] Han T., Guan Z., Chi M., Hu B., Li T., Zhang X., 2017, "Multi-formation control of nonlinear leader-following multi-agent systems," ISA Transactions, DOI: 10.1016/j.isatra.2017.05.003.
[8] Semsar-Kazerooni E., Khorasani K., 2009, "Multi-agent team cooperation: A game theory approach," Automatica, 45, 2205-2213.
[9] Mao D., He Y., Ye X., Yu M., 2011, "Inverse optimal stabilization of cooperative control in networked multi-agent systems," Control and Decision Conference (CCDC), 1031 - 1037.
[10] Tijs S., "Introduction to Game Theory," India: Hindustan Book Agency, 2003.
[11] Isaacs R., "Differential Games," New York, Wiley, 1965.
[12] Tolwinski B., Havrie A., Leimann G., 1986, "Cooperative equilibrium in differential games," Journal of Mathematical Analysis and Applications, 119, 182-202.
[13] Esparza L. G., Torres G. M., Saynes Torres L. M., 2013, "A brief introduction to differential games," International Journal of Physical and Mathematical Sciences, 4(1), 396-411.
[14] Başar T., Olsder G. "Dynamic Non-cooperative Game Theory," 2nd edition, Classics in Applied Mathematics. SIAM: Philadelphia, 1999.
[15] Freiling G., Jank G., Abou-Kandil H., 2002, "On global existence of Solutions to Coupled Matrix Riccati equations in closed loop Nash Games," IEEE Transactions on Automatic Control, 41(2), 264- 269.
[16] Gajic Z., Li T., 1988, "Simulation results for two new algorithms for solving coupled algebraic Riccati equations," Third Int. Symp. On Differential Games. Sophia, Antipolis, France.
[17] Sutton R., Barto A., "Reinforcement Learning–An Introduction," Massachusetts: Cambridge, MIT Press, 1998.
[18] Werbos P., "Approximate dynamic programming for real-time control and neural modeling Handbook of Intelligent Control," Ed. D.A. White and D.A. Sofge, New York: Van Nostrand Reinhold, 1992.
[19] Vrabie D., Pastravanu O., Lewis F. L., Abu-Khalaf M., 2009, "Adaptive optimal control for continuous-time linear systems based on policy iteration," Automatica, 45(2), 477–484.
[20] Vamvoudakis K., Lewis F.L., 2011, "Multi-player non-zero-sum games: online adaptive learning solution of coupled Hamilton-Jacobi equations," Automatica, 47, 1556-1559.
[21] Vrabie D., Lewis F., 2010, "Integral Reinforcement Learning for Online Computation of Feedback Nash Strategies of Nonzero-Sum Differential Games," 49th IEEE Conference on Decision and Control, Atlanta, GA, USA.
[22] Vrabie D., Lewis F.L., 2011, "Integral reinforcement learning for finding online the feedback Nash equilibrium of Nonzero-sum differential games," Advances in Reinforcement Learning, Intech, 2011.
[23] Vamvoudakis K. G., Lewis F. L., Hudas G. R., 2012, "Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality," Automatica, 48, 1598–1611.
[24] Abouheaf M. I., Lewis F. L., 2013, Multi-Agent Differential Graphical Games: Nash Online Adaptive Learning Solutions, 52nd IEEE Conference on Decision and Control, Florence, Italy.
[25] Tatari F., Naghibi-Sistani M. B., Vamvoudakis K. G., 2017, "Distributed Optimal Synchronization Control of Linear Networked Systems under Unknown Dynamics," Proc. American Control Conference, 668-673, Seattle, WA.
[26] Tatari F., Naghibi-S M., 2015, "Distributed Optimal Control of Nonlinear Differential Graphical Games based on Reinforcement Learning," Journal of Control, 8 (4),15-30.
[27] J. Li, H. Modares, T.Chai, F. L. Lewis, L. Xie, 2017, " Off-policy reinforcement learning for synchronization in multiagent graphical games," IEEE transactions on neural networks and learning systems, 28(10), 2434 - 2445.
[28] Kyriakos G. Vamvoudakis, 2017 "Q-learning for continuous-time graphical games on large networks with completely unknown linear system dynamics," International Journal of Robust and Nonlinear Control, 27(16), 2900-2920.
[29] Vamvoudakis K., Lewis F. L., 2011, "Online actor-critic algorithm to solve continuous-time infinite horizon optimal control problem," Automatica, 46, 787-788.
[30] Modares H., Lewis F.L., Naghibi-Sistani M.B., 2013, "Adaptive Optimal Control of Unknown Constrained-Input Systems Using Policy Iteration and Neural Networks," IEEE Transactions on neural networks and learning systems, 24(10), 1513-1525.
[31] Zhang H., Cui L., Luo Y., 2013, "Near-optimal control for nonzero-sum differential games of continuous-time nonlinear systems using single-network ADP," IEEE Trans. Cybern, 43, 206-216.
[32] Abu-Khalaf M., Lewis F. L., 2005, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," Automatica, 41, 779–791.
[33] Finlayson B. A., "The method of weighted residuals and variational principles," New York: Academic Press, 1990.

[34] Hornik K., Stinchcombe M., White H., 1990, "Universal approximation of an unknown mapping and its derivatives using multi layer feedforward networks," Neural Networks, 3(5), 551–560.
[35] Hardy G., Littlewood J., Polya G., "Inequalities," 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1989.
[36] Khalil H. K., "Nonlinear systems," Prentice-Hall, 1996.

[DOR: 20.1001.1.20088345.1397.12.3.6.4]
[DOI: 10.29252/joc.12.3.13]