

مقالات بخش فارسی

- ۱ ارائه روش مقاوم جهت ترکیب داده در سیستم تلفیقی و مقایسه آن با فیلتر کالمن توسعه یافته در شرایط بحرانی GPS/SDINS
علی اسدیان، بهزاد مشیری، علی خاکی صدیق
- ۱۲ طراحی H_{∞} مطمئن کنترل کننده برای سیستم های خطی تأخیردار با وجود نامعینی در تمام ماتریس های سیستم
علی ابوطالبی، سعید حسین نیا، فرید شیخ الاسلام
- ۲۰ ارائه روشی جدید در طراحی کنترلگر تطبیقی مقاوم برای سیستم های تأخیردار
سید کمال حسینی ثانی، حمیدرضا مؤمنی
- ۲۸ تأثیر بردار برتر در دینامیک فرآیند های QRDS و کاربرد آن در کنترل فرآیندهای دارای تأخیر زمانی
منصور شیروانی، منصوره اسماعیلی

مقالات بخش انگلیسی

- The Application of Non-Cooperative Game Theory in B2B Electronic Market** 1
Sahar Sohrabi, Mohammad Jafar Tarokh
- New Steady State Kalman Filter for Tracking High Maneuvering Targets** 8
Seyyed-Amin Hosseini, Mohammad Farrokhi
- A Behavioral Approach to Estimate Area of Expertise for Cooperative Q-Learning: A Self Referenced Strategy** 25
Ahmad Imanipour, Majid Nili Ahmadabadi, Babak N. Araabi, Masoud Asadpour, Roland Siegwar, Aram Zaerpour



نشریه علمی - پژوهشی، انجمن مهندسان کنترل و ابزار دقیق ایران، جلد ۲، شماره ۲، زمستان ۱۳۸۷

صاحب امتیاز: انجمن مهندسان کنترل و ابزار دقیق ایران

مدیر مسئول: پروفسور ایرج گودرزنیا

سردبیر: پروفسور علی خاکی صدیق

تلفن: ۸۴۰۶۲۳۱۷

ایمیل: sedigh@kntu.ac.ir

آدرس محل کار: خیابان دکتر شریعتی، پل سیدخندان، دانشکده برق دانشگاه صنعتی خواجه نصیرالدین طوسی

سمت: استاد دانشگاه

شورای سردبیری: پروفسور علی خاکی صدیق، دکتر حمید خالوزاده، دکتر علیرضا فاتحی

دبیر اجرایی: دکتر حمید خالوزاده

هیأت تحریریه:

پروفسور علی خاکی صدیق (استاد)، دکتر حمید خالوزاده (دانشیار)، پروفسور ایرج گودرزنیا (استاد)، پروفسور پرویز جبه دار مارالانی (استاد)، پروفسور علی غفاری (استاد)، دکتر هوشنگ حسیبی (استادیار)، دکتر محمدرضا جاهد مطلق (دانشیار)، دکتر کامبیز بدیع (دانشیار)، پروفسور رجب اصغریان (استاد)، دکتر حمیدرضا مومنی (دانشیار)، پروفسور علی وحیدیان کامیاد (استاد)، پروفسور سهراب خانمحمدی (استاد)، پروفسور سید کمال الدین نیکروش (استاد)، پروفسور مسعود شفیعی (استاد)، دکتر بتول لیبی (استادیار)، دکتر علیرضا فاتحی (استادیار)، پروفسور بهزاد مشیری (استاد)، پروفسور سراج الدین کاتبی (استاد)، مهندس امیر مقصودی پور (کارشناس ارشد)، مهندس کاووس فلامکی (کارشناس ارشد)، مهندس مهدی برادران مظفری (کارشناس ارشد)، مهندس غلامعباس رضانی (کارشناس ارشد).

هیأت مشاوران:

دکتر حمیدرضا مومنی، پروفسور علی غفاری، دکتر علی اکبر قره ویسی، دکتر محمد توکلی بینا، دکتر حمیدرضا تقی راد، دکتر محمد بطحایی، دکتر محمد تقی بهشتی، پروفسور بهزاد مشیری، پروفسور مسعود شفیعی، پروفسور رجب اصغریان، پروفسور علی خاکی صدیق، دکتر رضا کاظمی، دکتر سید علی اکبر موسویان، دکتر امیرحسین مرکزی دوابی، پروفسور محمد حایری، دکتر علیرضا خلیلی تهرانی، پروفسور حسین سیفی، دکتر احد کاظمی، دکتر حمید خالوزاده، دکتر علیرضا فاتحی، دکتر محمدرضا اکبرزاده توتونچی، دکتر میرعبادینی، دکتر حسین پدram، دکتر علی هارون آبادی، پروفسور علی وحیدیان کامیاد، دکتر جعفر حیرانی نوبی، دکتر بتول لیبی، پروفسور فرامرز حسین بابایی، دکتر پاکنوش کریم آقایی، دکتر بیژن معاونی، دکتر مهدی علیاری شوره دلی، دکتر محمد عاروان

هیأت مدیره انجمن مهندسان کنترل و ابزار دقیق:

مهندس عباس شعری مقدم، دکتر کیوان مسروری، دکتر حمیدرضا مومنی، پروفسور بهزاد مشیری، فرزاد جعفری کاظمی، دکتر حمید خالوزاده، مهندس علیرضا رستگاری

مقالات بخش فارسی

- ۱ ارائه روشی مقاوم جهت ترکیب داده در سیستم تلفیقی و مقایسه آن با فیلتر کالمن توسعه یافته در شرایط
بحرانی GPS/SDINS

علی اسدیان، بهزاد مشیری، علی خاکی صدیق

- ۱۲ طراحی مطمئن کنترل کننده H_{∞} برای سیستم های خطی تأخیر دار با وجود نامعینی در تمام ماتریس
های سیستم

علی ابوطالبی، سعید حسین نیا، فرید شیخ الاسلام

- ۲۰ ارائه روشی جدید در طراحی کنترلگر تطبیقی مقاوم برای سیستم های تاخیر دار

سید کمال حسینی ثانی، حمیدرضا مومنی

- ۲۸ تاثیر بردار برتر در دینامیک فرآیندهای QRDS و کاربرد آن در کنترل فرآیندهای دارای تاخیر
زمانی

منصور شیروانی، منصوره اسماعیلی

مقالات بخش انگلیسی

- The application of non-cooperative Game theory in B2B electronic market** 1

Sahar Sohrabi, Mohammad Jafar Tarokh

- New Steady State Kalman Filter for Tracking High Maneuvering Targets** 8

Seyyed-Amin Hosseini, Mohammad Farrokhi

- A Behavioral Approach to Estimate Area of Expertise for Cooperative Q-Learning:
A Self Referenced Strategy** 25

Ahmad Imanipour, Majid Nili Ahmadabadi, Babak N. Araabi, Masoud Asadpour, Roland Siegwär, Aram Zaerpour

مجله کنترل، مجله ای علمی - پژوهشی است که در برگیرنده تازه ترین نتایج تحقیقات نظری و کاربردی در علوم مختلف مرتبط با مهندسی کنترل و ابزار دقیق میباشد. از میان مباحث مورد نظر این مجله میتوان به موارد زیر اشاره نمود:

- ۱) مدل سازی، شناسایی، شبیه سازی و بهینه سازی سیستمها.
- ۲) تحلیل و طراحی سیستم های کنترل پیشرفته همچون سیستم های کنترل خطی و غیرخطی، سیستمهای کنترل تطبیقی، کنترل مقاوم و کنترل بهینه، سیستمهای کنترل هوشمند، سیستمهای کنترل تصادفی، سیستمهای کنترل گسسته پیشامد و ترکیبی، سیستمهای ابعاد وسیع.
- ۳) مکترونیک و رباتیک.
- ۴) ابزار دقیق و سیستمهای ترکیب داده و اطلاعات سنسوری.
- ۵) اتوماسیون صنعتی همچون سیستمهای کنترل گسترده، رابط انسان - ماشین، سیستمهای ایمنی و تشخیص خطا، سیستمهای زمان حقیقی و سیستمهای کنترل سوپروایزری.

کاربردهای مورد علاقه این مجله، وسیع بوده و می تواند در برگیرنده انواع فرآیندها باشد. از آن میان می توان به موارد زیر اشاره نمود:

- ۱) سیستمهای هدایت و ناوبری
- ۲) فرآیندهای صنعتی شامل فرآیندهای شیمیایی و بیوتکنولوژی.
- ۳) فرآیندهای استخراج و فراوری موارد معدنی.
- ۴) سیستمهای حمل و نقل و خودروهای خودکار.
- ۵) تولید و توزیع نیروی برق.
- ۶) مهندسی محیط زیست و هواشناسی.
- ۷) مهندسی تکنولوژی تولید.
- ۸) سیستمهای اقتصادی و مالی.
- ۹) سیستمهای اطلاعاتی، مخابراتی و شبکه های صنعتی.
- ۱۰) مهندسی پزشکی.
- ۱۱) سیستمهای آموزش هوشمند.

از کلیه پژوهشگران و کارشناسان فعال در زمینه های مرتبط با مهندسی کنترل و ابزار دقیق دعوت بعمل می آید تا مقالات و نتایج آخرین دستاوردهای علمی و پژوهشی خود را به این مجله ارسال نمایند. خواهشمند است مقالات خود را به صورت الکترونیکی به آدرس control@isice.ir ارسال فرمایید. برای کسب اطلاعات بیشتر و دریافت نحوه تهیه و ارسال مقالات می توانید به سایت مجله با آدرس www.isice.ir مراجعه نمایید.

ارائه روشی مقاوم جهت ترکیب داده در سیستم تلفیقی GPS/SDINS

و مقایسه آن با فیلتر کالمن توسعه یافته در شرایط بحرانی

علی اسدیان^۱، بهزاد مشیری^۲، علی خاکی صدیق^۳

^۱ فارغ‌التحصیل دانشکده برق و کامپیوتر، گروه کنترل، دانشگاه تهران، a.asadian@ece.ut.ac.ir

^۲ قطب علمی کنترل و پردازش هوشمند، دانشکده برق و کامپیوتر، دانشگاه تهران، moshiri@ut.ac.ir

^۳ دانشکده مهندسی برق، دانشگاه خواجه نصیرالدین طوسی تهران، sedigh@eedt.kntu.ac.ir

چکیده: در این مقاله روشی مقاوم و مبتنی بر فیلتر ذره‌ای برای مساله تخمین حالت در سیستم تلفیقی GPS/SDINS ارائه می‌شود. روش جاافتاده و کلاسیکی چون فیلتر کالمن توسعه یافته که مبتنی بر خطی‌سازی مدل می‌باشد، در ناوبری زمینی در موارد بحرانی مانند محیط‌های شهری که کاهش تعداد ماهواره‌های تحت رویت از حد مجاز لازم بدلیل موانع فیزیکی مختلف کاملاً محتمل است، به سرعت واگرا می‌شود، درحالی‌که با بکارگیری روش معرفی شده در این مقاله می‌توان ناوبری را با دقت نسبتاً مطلوبی در شرایط نامساعد رویت انجام داد.

کلمات کلیدی: فیلتر ذره‌ای، فیلتر کالمن توسعه یافته، ناوبری زمینی، قطع ارتباط ماهواره‌ای، سیستم تلفیقی GPS/SDINS

Abstract: In this paper, a robust state space estimation technique based on particle Kalman filters is presented for GPS/SDINS. The well known extended Kalman filters can fail under critical condition such as loss of satellite data in urban environment. This is due to physical obstacles present in such environments and leads to Kalman filter divergence. It is shown in this paper that with the proposed technique, this problem is mainly overcome.

Keywords: Particle Filter, Extended Kalman Filter,

کنترل و تئوری تخمین و برخی دیگر نیز بر پایه روش‌های نوین هوشمند پایه‌گذاری شده‌اند [۲].

۱- مقدمه

بدیهی است که فیلتر کالمن برای سیستم‌های خطی با نویز گوسی جواب بهینه است، ولی در غیر اینصورت فیلتر کالمن توسعه یافته بعنوان جایگزینی مناسب برای فیلتر کالمن استاندارد، روشی کلاسیک در حل چنین مساله ایست. در [۷-۱۰] در این زمینه بررسی‌هایی صورت گرفته است. جدا از مساله زیربهنگی^۲ فیلتر کالمن توسعه یافته در یک سیستم غیرخطی و از جمله سیستم تلفیقی GPS/SDINS، معضل قطع ارتباط ماهواره‌ای در ناوبری زمینی و در کاربردهای شهری، امری اجتناب ناپذیر است که از نقاط ضعف فیلتر کالمن محسوب می‌شود [۱۶].

امروزه سیستم تلفیقی GPS/SDINS به دلیل بالا بودن دقت و قابلیت اعتماد کاربرد وسیعی در سیستم‌های ناوبری پیدا کرده است [۱]. منتهی غیرخطی بودن مدل دینامیکی سیستم فوق‌الذکر بهمراه پیچیدگی‌های جانبی آن از جمله وجود منابع مختلف خطا، مساله تخمین حالت و ناوبری دقیق را دچار مشکل می‌کند. آنطور که از مطالعه منابع مختلف برآمد، رایج‌ترین روش جهت ترکیب اطلاعات در سیستم ناوبری، استفاده از فیلتر کالمن و مشتقات آن مانند فیلتر کالمن توسعه یافته می‌باشد. بر این اساس روش فوق به عنوان رویکردی متداول و یک محک^۱ کلی جهت مقایسه عملکرد سایر روش‌ها، مطرح می‌باشد. برخی از این روش‌های ترکیب داده بر اساس تئوری‌های کلاسیک علم

² Suboptimality

¹ Benchmark

تشعشع، تاخیر آتمسفریک، خطای چند مسیری، خطای ناشی از اختلاف ساعت و همچنین نویز گیرنده، چند عامل موثر در کاهش دقت این سیستم ناوبری می‌باشد. در عوض در حالت عادی خطای آن حتی برای زمان‌های طولانی محدود است [۳]. در نتیجه به منظور افزایش صحت ناوبری و قابلیت اعتماد سیستم و رفع معایب هر سیستم توسط دیگری، استفاده از هر دو نوع سیستم ناوبری همراه با یک الگوریتم مناسب ترکیب توصیه می‌شود.

در [۳-۶] به مدل‌های دینامیکی جامعی برای سیستم فوق اشاره شده است. معادلات کلی حاکم بر یک SDINS به فرم زیر است:

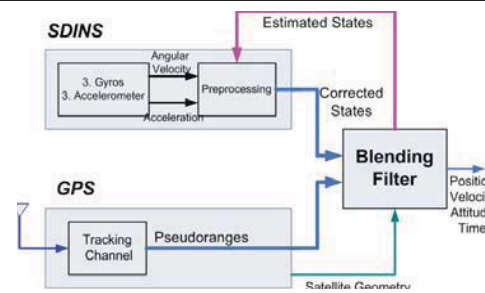
$$\begin{pmatrix} \dot{r}^n \\ \dot{v}^n \\ \dot{C}_b^n \end{pmatrix} = \begin{pmatrix} v^n \\ C_b^n f^b - (2\Omega_{ie}^n + \Omega_{en}^n)v^n + g^n \\ C_b^n (\Omega_{ib}^b - \Omega_{en}^b) \end{pmatrix} \quad (1)$$

در رابطه بالا، r^n و v^n بترتیب بردار مکان و سرعت در دستگاه ناوبری و C_b^n ماتریس تبدیل از دستگاه بدنی به ناوبری می‌باشند. f^b نیز بردار نیرو در مختصات بدنی و g^n شتاب جاذبه است. ماتریس $\Omega_{km}^l = \Omega(\omega_{km}^l)$ نیز سرعت نسبی دستگاه m نسبت به دستگاه k است که در چارچوب مختصاتی دستگاه l بیان می‌شود. ω_{ie}^n سرعت زاویه‌ای زمین نسبت به دستگاه اینرسی در دستگاه ناوبری می‌باشد. خروجی سیستم اندازه‌گیر اینرشیا که مقادیر شتاب‌ها و سرعت‌های زاویه‌ای می‌باشند، در حکم ورودی برای مجموعه معادلات بالا می‌باشند. پس حل مساله ناوبری اینرشیا به معنی حل معادلات دیفرانسیل کلی (۱) با فرض داشتن مقادیر لحظه‌ای f^b و ω_{ib}^b و همچنین یک شرط اولیه برای مساله می‌باشد.

آنطور که در مراجع مختلف مشاهده شد، بدلیل پیچیدگی نسبی مدل واقعی سیستم تلفیقی، عموماً مدل‌های دینامیکی ساده‌ای برای پیاده‌سازی الگوریتم ترکیب بکار می‌رود [۹-۱۰]، که این امر تاحدی جامعیت روش‌های ارائه شده را زیر سوال می‌برد لذا در این مقاله سعی شد تا حد امکان مدل کاملی را منظور کرد. در نهایت از فرم باز شده معادله (۱)، به شرح زیر استفاده شد [۱۴] و [۱۶].

جدول ۱: پارامترهای بیضی گون مینا در دستگاه مختصاتی WGS84

مقدار	نام پارامتر
۶۳۷۸/۱۳۷ کیلومتر	a (نصف قطر بزرگ)
۶۳۵۶/۷۵۲ کیلومتر	b (نصف قطر کوچک)
$10^{-3} \times 7/292115$ رادیان بر ثانیه	ω_{ie} (سرعت زاویه‌ای)



شکل ۱: بلوک دیاگرام کلی ترکیب داده در سیستم تلفیقی GPS/SDINS

در این مقاله روشی نوین برای بهینه‌سازی و تخمین در مواردی که با سیستم‌هایی غیر خطی و یا نویز غیر گوسی مواجهیم، ارائه می‌شود. ایده اصلی این روش استفاده از نظریه فیلترهای ذره‌ای ۳ و بدست آوردن تابع چگالی احتمال مشروط بردار حالت با توجه به بردار اندازه‌گیری‌ها و استفاده از تئوری بیز می‌باشد. در نتیجه می‌توان تخمین بهینه‌ای از هر تابع دلخواه از عناصر بردار حالت بر اساس معیار حداقل بودن واریانس خطا بدست آورد. بلوک دیاگرام کلی چنین سیستمی در حالت ترکیب داده متمرکز ۴ در شکل (۱) نمایش داده شده است که اساس کار ما نیز قرار خواهد گرفت. بخش مهم این دیاگرام، بلوک ترکیب و یا تخمین‌گر می‌باشد که می‌تواند هریک از دو فیلتر فوق‌الذکر فرض شود که موضوع اصلی این فعالیت پژوهشی می‌باشد.

۲- معرفی سیستم تلفیقی GPS/SDINS

سیستم‌های ناوبری به دو دسته کلی مکان‌یاب مانند سیستم ماهواره‌ای GPS و تعیین موقعیت اینرسی یا INS تقسیم می‌شوند. سیستم INS از جمله دقیق‌ترین و امن‌ترین سیستم‌های ناوبری است چراکه بدون هیچ نوع دریافت و ارسال سیگنال مخابراتی تنها با اندازه‌گیری نیروی وارد بر جسم و سرعت زاویه‌ای مربوطه، موقعیت و وضعیت را در هر لحظه تعیین می‌کند. عیب آن نیز این است که خطای ناوبری آن با گذشت زمان با توجه به بایاس اولیه بطور تصاعدی افزایش می‌یابد و در زمان‌های طولانی موجب انحراف قابل توجه جسم متحرک می‌شود.

از طرفی سیستم GPS به علت ارتباط مخابراتی با دنیای خارج، براحتی می‌تواند توسط گردانندگان سیستم مخابراتی مورد اغتشاش واقع شود. پهنای باند آن پایین بوده و مسائلی چون تداخلات ناشی از

³ Particle Filters (PFs)

⁴ Centralized Data Fusion

⁵ World Geodetic System

با صرف نظر از اثر تغییرات شتاب جاذبه زمین، در نهایت مدل دینامیکی INS توسط روابط (۷-۱۰) بیان می‌شود. در این روابط g شتاب جاذبه، $(\tilde{a}_u \tilde{a}_v \tilde{a}_w)^T$ مقادیر اندازه‌گیری شده توسط سه شتاب‌سنج و $(b_u b_v b_w)^T$ بایاس اولیه آن می‌باشد. $W_t^{b_{acc}}$ و $W_t^{b_{gyro}}$ نیز فرآیندهای تصادفی وینر با متوسط صفر و واریانس مناسب می‌باشند.

$$d \begin{pmatrix} \phi \\ \lambda \\ h \end{pmatrix} = \begin{pmatrix} \frac{1}{R_\phi + h} & 0 & 0 \\ 0 & \frac{1}{(R_\lambda + h) \cos(\phi)} & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} V_N \\ V_E \\ V_D \end{pmatrix} dt$$

$$d \begin{pmatrix} V_N \\ V_E \\ V_D \end{pmatrix} = \begin{pmatrix} \frac{-V_E^2 \tan(\phi)}{R_\lambda + h} - 2\omega_{ie} \sin(\phi) V_E + \frac{V_N V_D}{R_\phi + h} \\ \frac{V_E V_N \tan(\phi)}{R_\phi + h} + \omega_{ie} \{\sin(\phi) V_N + 2 \cos(\phi) V_D\} + \frac{V_E V_D}{R_\lambda + h} \\ \frac{-V_N^2}{R_\phi + h} - 2\omega_{ie} \cos(\phi) V_E - \frac{V_E^2}{R_\lambda + h} \end{pmatrix} dt$$

$$+ R_{b2g} \left\{ \begin{pmatrix} \tilde{a}_u \\ \tilde{a}_v \\ \tilde{a}_w \end{pmatrix} + \begin{pmatrix} b_u \\ b_v \\ b_w \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ g \end{pmatrix} \right\} dt + dW_t^v$$

$$d \begin{pmatrix} b_u \\ b_v \\ b_w \end{pmatrix} = -a_{acc} \begin{pmatrix} b_u \\ b_v \\ b_w \end{pmatrix} dt + dW_t^{b_{acc}}$$

$$d \begin{pmatrix} b_p \\ b_q \\ b_r \end{pmatrix} = -a_{gyro} \begin{pmatrix} b_p \\ b_q \\ b_r \end{pmatrix} dt + dW_t^{b_{gyro}} \quad (7-10)$$

مدلسازی خطای ساعت: مدل مربوط به خطای ساعت گیرنده یعنی δ_t به صورت انتگرال یک فرآیند تصادفی x_t بیان می‌شود. W_t^x نیز یک فرآیند وینر با متوسط صفر و واریانس σ_w^2 می‌باشد.

$$dx_t = -a_{clock} x_t dt + dW_t^x \quad (11)$$

$$d\delta_t = x_t dt \quad (12)$$

مدلسازی تاخیر آتمسفریک: جهت مدلسازی خطای تروپوسفریک به عنوان منبع اصلی خطای سیستم GPS که قابل مدلسازی بوده و عموماً غیرقابل حذف می‌باشد، از مدل ساده شده گود-گودمن^۸ و معادله (۱۳)

	دوران زمین)
$\frac{a-b}{a} = 3.35281113 \times 10^{-5}$	f (میزان تخت بودن)
$\sqrt{f(1-f)} = 0.05781$	e (خروج از مرکز)

$$\begin{pmatrix} V_N \\ V_E \\ V_D \end{pmatrix} = \begin{pmatrix} (R_\phi + h)\dot{\phi} \\ (R_\lambda + h)\cos(\phi)\dot{\lambda} \\ -\dot{h} \end{pmatrix} \quad (2)$$

$$R_\lambda = \frac{a(1-e^2)}{\sqrt{\{1-e^2 \sin^2(\phi)\}^3}} \quad (3)$$

$$R_\phi = \frac{a}{\sqrt{1-e^2 \sin^2(\phi)}} \quad (4)$$

عموماً مکان جسم بر حسب پارامترهای λ (طول جغرافیایی)، ϕ (عرض جغرافیایی) و h (ارتفاع) بیان می‌شود. پارامترهای بیضی‌گون مینا جهت مدل کردن کره زمین در جدول (۱) آمده است. سرعت جسم متحرک نیز معمولاً در دستگاه مختصات محلی NED^۶ بیان می‌شود. رابطه بین بردار سرعت و پارامترهای جغرافیایی به صورت روابط (۲-۴) است. INS موردنظر در اینجا از نوع SDINS یعنی بدون صفحه پایدارکننده^۷ است و اندازه‌گیری‌های سه شتاب‌سنج و سه ژيروسکوپ نسبت به بدنه واحد INS انجام می‌شود. ماتریس تبدیل بین این مختصات بدنه و مختصات محلی NED، در معادله دیرفرانسیل (۵) صدق می‌کند:

$$dR_{b2g} = R_{b2g} \begin{pmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{pmatrix} dt \quad (5)$$

$$\omega_{gb}^b = \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} \tilde{p} \\ \tilde{q} \\ \tilde{r} \end{pmatrix} + \begin{pmatrix} b_p \\ b_q \\ b_r \end{pmatrix} - R_{g2b} \begin{pmatrix} \omega_{ie} \cos(\phi) + \frac{V_E}{R_\lambda + h} \\ -\frac{V_N}{R_\phi + h} \\ -\omega_{ie} \sin(\phi) + \frac{V_E \tan(\phi)}{R_\lambda + h} \end{pmatrix} \quad (6)$$

بردار سرعت زاویه‌ای جسم متحرک در مختصات بدنه نیز یعنی: $\omega_{gb}^b = (p \ q \ r)^T$ در معادله (۶) صدق می‌کند که در این معادله $(\tilde{p} \ \tilde{q} \ \tilde{r})^T$ مقادیر اندازه‌گیری شده توسط ژيروسکوپ‌ها و $(b_p \ b_q \ b_r)^T$ بایاس اولیه مربوطه می‌باشد.

^۸ Goad & Goodman

^۶ North East Down
^۷ Strapdown INS

مختصاتی^۹ ECEF می‌باشد. جهت حل مساله ناهمزمانی داده‌های دو سیستم ناوبری نیز از یک درون‌یابی خطی ساده استفاده شد [۸] و [۱۶]. توجه شود عموماً برای تخمین بردار حالت بر اساس اندازه‌گیری شبه‌فاصله‌ها با کمک روشی غیرهوشمند، در هر لحظه حداقل باید ۴ ماهواره قابل رویت باشد. برای افزایش دقت ناوبری نیز باید بیش از این تعداد یعنی مثلاً ۵ ماهواره در هر لحظه برای گیرنده قابل رویت باشد. در ناوبری شهری در حالات بحرانی، موانع طبیعی و یا غیر طبیعی ممکن است مانع از رویت مناسب و حتی کاهش تعداد ماهواره‌های تحت رویت شود. این امر متناظر با عبور وسیله متحرک از جلوی ساختمانهای بلند، زیر پل‌ها و یا داخل تونل‌ها و یا اصولاً وجود هر نوع مانع فیزیکی که موجب تضعیف و یا قطع ارتباط مخابراتی گیرنده GPS نصب شده بر روی وسیله نقلیه با هریک از ماهواره‌ها گردد، می‌باشد.

این قطع ارتباط ماهواره‌ای طبق [۸]، منجر به میل کردن فاکتورهای هندسی تعیین موقعیت به سمت بی‌نهایت و واگرایی سریع فیلتر کالمن توسعه یافته خواهد شد. بدیهی می‌باشد که شاخص فوق یک پارامتر مهم جهت ارزیابی میزان کارایی الگوریتم ترکیب داده بشمار می‌رود و ارائه یک روش مقاوم نسبت به معضل فوق، در امر ناوبری زمینی یک نقطه قوت بسیار مهم بشمار می‌رود. بنابراین در این مقاله، شاخص خطای ناوبری در حالت قطع ارتباط ماهواره‌ای، به عنوان یک معیار اساسی جهت مقایسه دو روش مختلف مورد بحث در تلفیق دو سیستم ناوبری GPS و INS و تعیین میزان مقاومت هریک، مورد بررسی قرار خواهد گرفت. همانطور که نشان داده خواهد شد، در چنین مواقعی روش ارائه شده در [۱۵] می‌تواند راه حل مناسبی برای مساله تخمین حالت در مساله ناوبری باشد.

۳- فیلتر ذره‌ای و الگوریتم پیشنهادی جهت تخمین حالت

همانطور که ذکر شد استفاده از فیلتر کالمن توسعه یافته یک روش متداول در برخورد با مساله تخمین در سیستم‌های غیرخطی و بخصوص در کاربرد ناوبری است [۹-۶]. در اینجا روشی کارا برای حل مساله تخمین در مواجهه با سیستم‌های غیرخطی و غیرگوسی و مقاوم در برابر قطع ارتباط ماهواره‌ای در سیستم تلفیقی ارائه می‌شود. این روش جزء روش‌های آماری مونت کارلو بوده و بر پایه استفاده از قانون اعداد بزرگ است و تابع چگالی احتمال شرطی را به صورت مجموع

استفاده می‌کنیم که در آن θ زاویه فراز رویت ماهواره، R شعاع زمین و h_d و h_w به ترتیب ارتفاع لایه‌های تروپوسفریک مرطوب و خشک، T دما، p فشار محلی و pe فشار جزئی بخار آب می‌باشند [۱۶]. پارامترهای تقریبی این مدل نیز طبق جدول (۲) فرض می‌شوند.

$$\Delta_{tropo}(\theta) = 10^{-6} \{ N_d^{tropo} \{ \sqrt{(R+h_d)^2 - R^2 \cos^2(\theta)} - R \sin(\theta) \} + N_w^{tropo} \{ \sqrt{(R+h_w)^2 - R^2 \cos^2(\theta)} - R \sin(\theta) \} \} \quad (13)$$

جدول ۲: پارامترهای مدل خطای تروپوسفریک

نام پارامتر	مقدار
R	۶۳۶۷/۴۳ کیلومتر
h_w	۱۱۰۰۰ متر
h_d	$148/72(T-273/16)+40/136$
N_w^{tropo}	$\frac{pe}{T^2} + 710/471 - 12/96 \frac{pe}{T}$
N_d^{tropo}	$\frac{pe}{77/64 T}$

شبه‌فاصله‌ها و معادله مربوط به بردار اندازه‌گیری: در این مقاله ناوبری صرفاً بر اساس اندازه‌گیری شبه‌فاصله‌ها انجام می‌پذیرد. شبه‌فاصله‌ها بصورت فاصله نسبی بین هریک از ماهواره‌ها با جسم متحرک که توسط منابع خطا از مقدار واقعی منحرف می‌شوند، توسط روابط ۱۴ الی ۱۶ بیان می‌شود.

$$\rho_i = \sqrt{(X_{si} - X_m)^2 + (Y_{si} - Y_m)^2 + (Z_{si} - Z_m)^2} + c\delta_i + \Delta_{tropo}(\theta_i) + \eta_i$$

$$\theta_i = \frac{\pi}{2} - \text{Arc cos} \left\{ \frac{X_m(X_{si} - X_m) + Y_m(Y_{si} - Y_m) + Z_m(Z_{si} - Z_m)}{\sqrt{(X_{si} - X_m)^2 + (Y_{si} - Y_m)^2 + (Z_{si} - Z_m)^2}} \right\} \times \frac{1}{\sqrt{X_m^2 + Y_m^2 + Z_m^2}}$$

$$\begin{pmatrix} X_m \\ Y_m \\ Z_m \end{pmatrix} = \begin{pmatrix} (R_\phi + h) \cos(\lambda) \cos(\phi) \\ (R_\phi + h) \cos(\lambda) \sin(\phi) \\ (R_\phi(1 - e^2) + h) \sin(\lambda) \end{pmatrix}$$

(۱۴-۱۶)

در این روابط C سرعت نور، δ_i بایاس ساعت گیرنده، θ_i زاویه فراز رویت ماهواره نام، η_i نویز اندازه‌گیری، (X_{si}, Y_{si}, Z_{si}) مختصات ماهواره نام و (X_m, Y_m, Z_m) مختصات گیرنده در چارچوب

⁹ Earth Centered Earth Fixed

\hat{B}_w معکوس مور-پنرس ماتریس B_w می‌باشد. اکنون می‌خواهیم $P(x_{k+1} | y_{1:k+1})$ را بدست آوریم. فرض می‌کنیم y_{k+1} از طریق اندازه‌گیری در دسترس است. طبق قاعده بیز:

$$P(x_{k+1} | y_{1:k+1}) = \frac{P(y_{k+1} | x_{k+1})P(x_{k+1} | y_{1:k})}{P(y_{k+1} | y_{1:k})} \quad (23)$$

$P(x_{k+1} | y_{1:k})$ از طریق رابطه (۲۱) قبلاً حساب شد. $P(y_{k+1} | x_{k+1})$ هم از طریق مدل اندازه‌گیری مشابه رابطه (۲۲) توسط رابطه (۲۴) حساب می‌شود.

$$P(y_{k+1} | x_{k+1}) = P_{v_{k+1}}(z_{k+1} - h(x_{k+1})) \quad (24)$$

مخرج کسر هم طبق رابطه (۲۵) بدست می‌آید.

$$P(y_{k+1} | y_{1:k}) = \int P(y_{k+1} | x_{k+1})P(x_{k+1} | y_{1:k})dx_{k+1} \quad (25)$$

بنابراین در مرحله بروز کردن با توجه به روابط ذکر شده، $P(x_{k+1} | y_{1:k+1})$ محاسبه می‌شود. انجام این دو مرحله بصورت بازگشتی منجر به یافتن کلی ضرایب بیز و تابع مجهول $P(x_k | y_{1:k})$ و حل مساله تخمین می‌شود. توجه شود که جواب بدست آمده گرچه دقیق است و هیچ نوع تقریبی در آن بکار نرفته است، ولی بدلیل بعد نامحدود فیلتر تخمین‌گر، عملاً قابل پیاده‌سازی نیست. روش مورد استفاده جهت تقریب رابطه بدست آمده در تعداد گام محدود، استفاده از فیلتر ذره‌ای به شرح زیر است:

اگر نمونه‌های $\{x_{0:k}^i\}_{i=1}^N$ براساس تابع $q(x_{0:k} | y_{1:k})$ بازسازی شوند^{۱۱}، وزن‌های مربوطه برابر خواهد بود با [۱۱]:

$$\omega_k^i \propto \frac{P(x_{0:k}^i | y_{1:k}^i)}{q(x_{0:k}^i | y_{1:k}^i)} \quad (26)$$

اگر $q(\cdot)$ را بتوان طوری انتخاب کرد که رابطه (۲۷) برقرار باشد، در اینصورت تخمین موردنظر از تابع چگالی احتمال را می‌توان برحسب توابع $P(x_{0:k-1} | y_{1:k-1})$ ، $P(y_k | x_k)$ و $P(x_k | x_{k-1})$ بدست آورد.

$$q(x_{0:k} | y_{1:k}) = q(x_k | x_{0:k-1}, y_{1:k})q(x_{0:k-1} | y_{1:k-1}) \quad (27)$$

این موضوع را در ذیل نشان می‌دهیم:

وزن‌یافته تعدادی تابع گسسته تقریب می‌زند. در [۱۴-۱۱] و [۱۶] بطور مفصل در مورد جزئیات ریاضی موضوع و همگرایی این فیلتر بحث شده است. در اینجا صرفاً نتایج کلی مورد استفاده قرار می‌گیرد. ابتدا به ذکر مبانی ریاضی موضوع پرداخته می‌شود. در خیلی از موارد می‌توان مدل سیستم را به صورت زیر بازنویسی کرد:

$$x_{k+1} = f(x_k, u_k) + B_w w_k \quad (17)$$

$$z_k = h(x_k) + B_v v_k = h(x_k) + \tilde{v}_k \quad (18)$$

این امر صرفاً جهت سادگی نمایش روابط انجام می‌شود. در این صورت: $R_k^v = \tilde{g}_k(x_k)R_k^v \tilde{g}_k^T(x_k)$ که منظور از R_k^v ماتریس کوواریانس نویز است. فرض هم می‌شود که تابع توزیع احتمال فرآیندهای w_k و \tilde{v}_k به ترتیب P_{w_k} و P_{v_k} باشد. هدف پیدا کردن تخمین بهینه یک تابع دلخواه مانند g از عناصر بردار حالت با معیار حداقل بودن واریانس خطاست.

$$\hat{g}(x_k)_{Minimum-Variance} = E[g(x_k) | Y_k] \quad (19)$$

که $Y_k = \{y_1, y_2, \dots, y_k\}$ مجموعه اندازه‌گیری‌هاست. در صورتیکه بتوان تابع توزیع احتمال شرطی $P(x_k | y_{1:k})$ را حساب کرد، تخمین بهینه از طریق رابطه (۲۰) حساب می‌شود. در اینجا به دنبال یافتن روشی بازگشتی برای محاسبه تابع توزیع فوق با در نظر گرفتن بردار اندازه‌گیری شده در هر لحظه می‌باشیم. این تابع توزیع دربرگیرنده کلیه اطلاعات آماری فرآیند X_k که از طریق مشاهده Y_k قابل حصول است، می‌باشد.

$$E[g(x_k) | Y_k] = \int_{\mathcal{X}^n} g(x_k) dP(x_k | Y_k) \quad (20)$$

این روش شامل دو بخش اساسی: ۱- پیش‌بینی و ۲- بروز کردن می‌باشد. در مرحله پیش‌بینی نخست فرض می‌کنیم که تابع $P(x_k | y_{1:k})$ در دسترس است و هنوز مقدار y_{k+1} اندازه‌گیری نشده است. با توجه به رابطه (۱۷) و معادله چاپمن-کومولوگورف خواهیم داشت:

$$P(x_{k+1} | y_{1:k}) = \int P(x_{k+1} | x_k)P(x_k | y_{1:k})dx_k \quad (21)$$

که جمله اول عبارت تحت انتگرال با استفاده از مدل سیستم به این صورت بدست می‌آید:

$$P(x_{k+1} | x_k) = P_{w_k}(\hat{B}_w(x_{k+1} - f(x_k, u_k))) \quad (22)$$

¹⁰ Importance Density Function

¹¹ Importance Sampling

با جایگزینی $q(\cdot)$ بهینه در رابطه (۳۴)، قانون زیر بدست می آید:

$$\begin{aligned} \omega_k^i &\propto \omega_{k-1}^i P(y_k | x_{k-1}^i) \\ &= \omega_{k-1}^i \int P(y_k | x_k') P(x_k' | x_{k-1}^i) dx_k' \end{aligned} \quad (۳۷)$$

این انتخاب ما را دچار ۲ معضل اساسی می کند. یکی اینکه ما نیازمندیم که از تابع $P(x_k | x_{k-1}^i, y_k)$ نمونه برداری کنیم و دیگری اینکه نیاز به انتگرال گیری از حالات گذرای ذرات داریم. بنابراین در خیلی از موارد بکاربردن مستقیم رابطه (۳۷) برای بدست آوردن $q(\cdot)$ بهینه، چندان عملی نیست. در نهایت در اغلب موارد $q(\cdot)$ را به شکل (۳۸) انتخاب می کنند و رابطه متناظر برای بروز کردن وزن ها طبق (۳۹) بدست می آید.

$$q(x_k | x_{k-1}^i, y_k) = P(x_k | x_{k-1}^i) \quad (۳۸)$$

$$\omega_k^i \propto \omega_{k-1}^i P(y_k | x_k^i) \quad (۳۹)$$

در نهایت الگوریتم زیر جهت مساله تخمین ارائه می گردد:

۱- بردار تصادفی $\{\tilde{x}_0^1, \tilde{x}_0^2, \dots, \tilde{x}_0^N\}$ با توجه به توزیع مفروض $P(x_0 | y_0) = P(x_0)$ انتخاب می شوند. هر کدام از این N بردار یک ذره نامیده می شود.

۲- مجموعه $\{\tilde{x}_k^1, \tilde{x}_k^2, \dots, \tilde{x}_k^N\}$ به دینامیک (۱۷) اعمال و $\{\tilde{x}_{k+1}^1, \tilde{x}_{k+1}^2, \dots, \tilde{x}_{k+1}^N\}$ با توجه به آن بدست می آید.

۳- طبق [۱۴]، اگر توزیع \tilde{v}_k گوسی باشد، انتخاب بهینه وزن های $\{\omega_k^i\}_{i=1}^N$ بر اساس روابط زیر می باشد:

$$\omega_k^i = \frac{Z_k^i}{\sum_{j=1}^N Z_k^j} \quad (۴۰)$$

$$Z_k^i(x, y) = \frac{\exp\left\{-\frac{1}{2} \sum_{j=1}^k \gamma^{k-j} \|y_j - h(\tilde{x}_j^i)\|_{R^{\bar{v}}}^2\right\}}{\exp\left\{-\frac{1}{2} \sum_{j=1}^k \gamma^{k-j} \|y_j\|_{R^{\bar{v}}}^2\right\}} \quad (۴۱)$$

به طوریکه: $\|a\|_{R^{\bar{v}}}^2 = a^T [R^{\bar{v}}]^{-1} a$, $a \in \mathbb{R}^m$ و $\gamma \in (0, 1)$ یک ضریب فراموشی و m بعد بردار خروجی است.

۴- با توجه به محاسبه وزن ها، مقادیر تخمینی در گام k ام طبق روابط (۴۲-۴۳) محاسبه می شوند.

$$\hat{P}(x_k | y_{1:k}) = \sum_{i=1}^N \omega_k^i \delta(x_k - \tilde{x}_k^i) \quad (۴۲)$$

$$\begin{aligned} P(x_{0:k} | y_{1:k}) &= \frac{P(y_k | x_{0:k}, y_{1:k-1}) P(x_{0:k} | y_{1:k-1})}{P(y_k | y_{1:k-1})} \\ &= \frac{P(y_k | x_{0:k}, y_{1:k-1}) P(x_k | x_{0:k-1}, y_{1:k-1})}{P(y_k | y_{1:k-1})} \\ &\quad \times P(x_{0:k-1} | y_{1:k-1}) \\ \Rightarrow P(x_{0:k} | y_{1:k}) &= \frac{P(y_k | x_k) P(x_k | x_{k-1})}{P(y_k | y_{1:k-1})} \times P(x_{0:k-1} | y_{1:k-1}) \\ \Rightarrow P(x_{0:k} | y_{1:k}) &\propto \\ P(y_k | x_k) P(x_k | x_{k-1}) &\times P(x_{0:k-1} | y_{1:k-1}) \end{aligned} \quad (۲۸-۳۰)$$

با جایگزین کردن ۲ رابطه اخیر در معادله (۲۶)، فرم بازگشتی لازم برای محاسبه وزن ها را بدست می آوریم.

$$\begin{aligned} \omega_k^i &\propto \\ \frac{P(y_k | x_k^i) P(x_k^i | x_{k-1}^i) \times P(x_{0:k-1}^i | y_{1:k-1})}{q(x_k^i | x_{0:k-1}^i, y_{1:k}) q(x_{0:k-1}^i | y_{1:k-1})} \end{aligned} \quad (۳۱)$$

$$\Rightarrow \omega_k^i = \beta \omega_{k-1}^i \frac{P(y_k | x_k^i) P(x_k^i | x_{k-1}^i)}{q(x_k^i | x_{0:k-1}^i, y_{1:k})} \quad (۳۲)$$

به علاوه اگر فرض شود:

$$q(x_k | x_{0:k-1}, y_{1:k}) = q(x_k | x_{k-1}, y_k) \quad (۳۳)$$

آنگاه تابع $q(\cdot)$ تنها به مقادیر x_{k-1} و y_k بستگی پیدا می کند و در سیستم پردازشگر تنها کفایت که تنها تکامل x_k^i و نه $x_{0:k-1}^i$ و $y_{1:k-1}$ ذخیره شود ($i = 1, 2, \dots, N$). در نتیجه قانون اصلاح شده بروز کردن وزن ها به این فرم بازنویسی می شود:

$$\omega_k^i \propto \omega_{k-1}^i \frac{P(y_k | x_k^i) P(x_k^i | x_{k-1}^i)}{q(x_k^i | x_{k-1}^i, y_k)} \quad (۳۴)$$

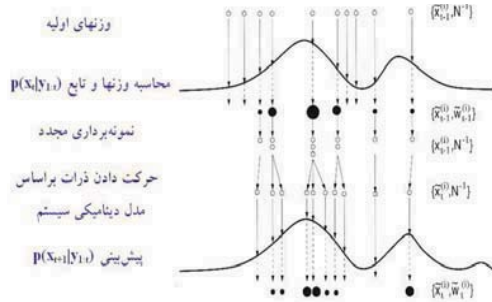
و تخمین موردنظر برابر خواهد بود با:

$$P(x_k | y_{1:k}) = \lim_{N \rightarrow \infty} \sum_{i=1}^N \omega_k^i \delta(x_k - x_k^i) \quad (۳۵)$$

تنها نکته مهم باقیمانده نحوه انتخاب تابع $q(\cdot)$ است که در اینجا به آن اشاره می شود. مهم ترین معیار برای انتخاب این تابع، حداقل شدن واریانس مجموعه وزن ها به منظور کاهش خطای تخمین و افزایش تعداد ذرات موثر می باشد. در [۱۱] این مقدار بهینه به این صورت معرفی شده است:

$$\begin{aligned} q(x_k | x_{k-1}^i, y_k)_{Optimal} &= P(x_k | x_{k-1}^i, y_k) \\ &= \frac{P(y_k | x_k, x_{k-1}^i) P(x_k | x_{k-1}^i)}{P(y_k | x_{k-1}^i)} \end{aligned} \quad (۳۶)$$

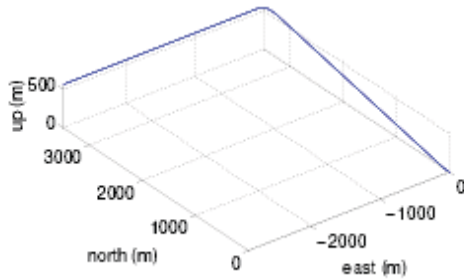
به شرط کافی بودن تعداد ذرات می‌باشد. در نهایت مراحل کلی الگوریتم فیلتر ذره‌ای را می‌توان در شکل (۲) مشاهده کرد.



شکل ۲: شمای کلی تخمین حالت توسط فیلتر ذره‌ای

۴- نتایج شبیه‌سازی

برای انجام شبیه‌سازی و کسب مقادیر عددی از نسخه آموزشی برنامه PROGEN و SatNav ارائه شده توسط GPSsoft استفاده شد [۱۷]. در این برنامه یک پروفایل حرکتی کند با دینامیک پایین و به طول ۱۱ دقیقه تولید و شبه‌فاصله‌ها، مقادیر خوانده شده توسط ژيروسکوپ‌ها و شتاب‌سنج‌ها و داده‌های مکانی مربوطه برای چهار ماهواره قابل رویت ارائه می‌شود. نمونه‌گیری واحد IMU با فرکانس ۱۰ Hz و خواندن مقدار شبه‌فاصله‌ها با فرکانس ۱ Hz صورت می‌پذیرد.



شکل ۳: پروفایل مسیر پرواز با دینامیک پایین

پارامترهای عددی مورد استفاده برای مدلسازی نیز به صورت زیر انتخاب شدند. این پارامترها یک واحد IMU ارزان قیمت با دقت متوسط را مدل می‌کند.

$$\hat{g}(x_k) = \sum_{i=1}^N \omega_k^i g(\tilde{x}_k^i) \quad (43)$$

۵- تعداد ذره‌های موثر حساب شود:

$$N_{eff} = \frac{1}{\sum_{i=1}^N (\omega_k^i)^2} \quad (44)$$

در صورتیکه رابطه: $(N_{th} = \frac{2N}{3})$ $N_{eff} \leq N_{th}$ برقرار باشد، الگوریتم نمونه‌گیری مجدد^{۱۲} اجرا شده و وزن‌ها به مقادیر:

$$\left\{ \omega_k^i \right\}_{i=1}^N = \frac{1}{N}$$

۶- با مراجعه به مرحله ۲ این روند تا رسیدن به دقت موردنظر مجدداً تکرار شود.

یکی دیگر از مسائل مهمی که به هنگام استفاده از این روش باید مورد توجه قرار گیرد، انتخاب بهینه N است. در [۱۱] ثابت شده است که انتخاب وزن‌ها طبق رابطه‌ای معادل با روابط (۴۰-۴۱)، موجب افزایش تعداد ذره‌های موثر و در نتیجه افزایش سرعت همگرایی می‌شود. هرچه تعداد ذره‌ها کمتر باشد، هزینه محاسباتی مساله کمتر است در عوض خطای تخمین بیشتر خواهد شد. اصولاً هر چه واریانس وزن‌های: $\left\{ \omega_k^i \right\}_{i=1}^N$ بیشتر باشد، تعداد موثر ذره‌ها کمتر می‌شود. در اینگونه موارد توسط الگوریتم نمونه‌گیری مجدد، ذره‌هایی که وزن متناظر با آنها بزرگ است هر یک به چند ذره با وزن نرمالیزه شده:

$$\left\{ \omega_k^i \right\} = \frac{1}{N}$$

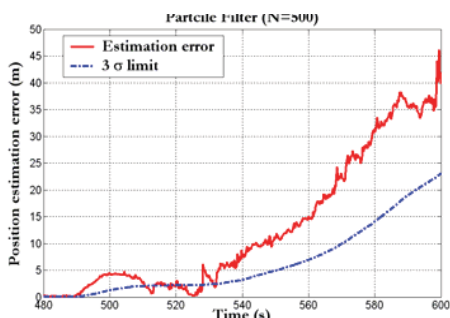
وزن متناظر با آنها کوچک است، تقریباً حذف می‌شوند. در عین حال تابع توزیع ناشی از ذرات قبل و بعد از این زیروال تقریباً ثابت خواهد ماند. برای نمونه‌گیری مجدد از روش ارائه شده در [۱۱] استفاده شد.

در صورت حذف مرحله ۵، پس از چند مرحله از بروز کردن وزن‌ها، نسبت وزن برخی از ذره‌ها نسبت به برخی دیگر بسیار کوچک شده و تدریجاً به سمت صفر میل می‌کند که این بدان معنی است که عملاً تعدادی از ذره‌ها تاثیر چندانی در معادلات (۴۲-۴۳) ندارند و هزینه محاسباتی بی‌بهره‌ای صرف بروز کردن وزن‌های آنها می‌شود (مشکل پراکندگی عددی وزن‌ها).

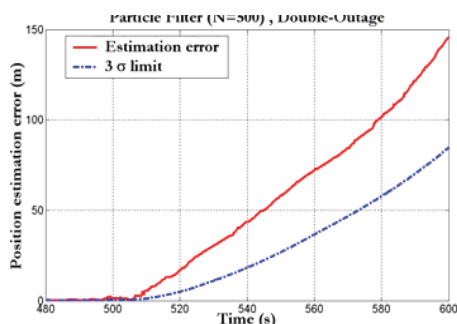
انتخاب ضریب فراموشی γ هم در مرحله ۳ جهت تضمین همگرایی یکنواخت $\hat{P}(x_k | y_{1:k})$ به مقدار واقعی آن با افزایش زمان

¹² Resampling Algorithm

شرایط بحرانی با دقت نسبتاً مطلوب مورد استفاده قرار گیرد. البته با کاهش تعداد ماهواره‌ها، طبق شکل (۶)، کم کم اثر ناپایداری در فیلتر ذره‌ای نیز مشاهده می‌شود.



شکل ۵: خطای تخمین مکان با فرض رویت ۳ ماهواره توسط فیلتر ذره‌ای به ازای $N=500$



شکل ۶: خطای تخمین مکان با فرض رویت ۲ ماهواره توسط فیلتر ذره‌ای به ازای $N=500$

نکته مهم در این دو شکل این است که در هر دو مورد در ابتدای بازه بحرانی، خطا در محدوده قابل قبولی قرار دارد که با توجه به اینکه عموماً موانع در ناوبری شهری، صرفاً در مقاطع زمانی کوتاهی ظاهر می‌شوند، روش فوق می‌تواند جایگزین خوبی برای روش متداولی چون فیلتر کالمن در چنین لحظاتی باشد. همین آزمایش به ازای تعداد ذرات ۲۰۰ و ۱۰۰۰ نیز انجام شد که نتایج در اشکال (۷) و (۸) مشاهده می‌شود.

$$a_{acc} = a_{gyro} = 0.0015$$

$$a_{clock} = 0.002$$

$$T = 300 \text{ K}$$

$$\sum w^{b_{acc}} = 4/9.5 \times 10^{-4} I$$

$$P = 1/0.13 \text{ bar}$$

$$\sum w^{b_{gyro}} = 0.009 I$$

$$pe = 35 \text{ mbar}$$

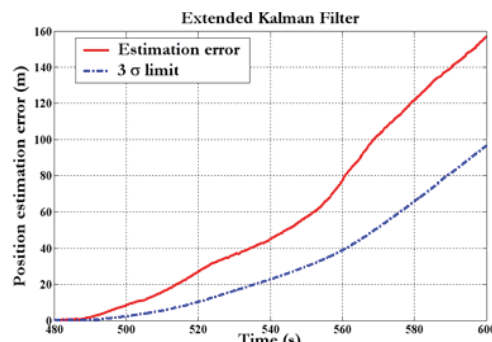
$$\sum w^{x} = 10^{-5} I$$

$$\sigma_{w^x} = 10^{-12} I$$

تعداد ذره‌ها هم با سعی و خطا ۵۰۰ انتخاب شد که توجه به پروفایل جسم متحرک که فاقد تغییرات خیلی سریع است، این تعداد کافی بنظر می‌رسد. البته جهت مقایسه عددی تعداد ذرات دیگری هم بررسی شد که در ادامه به آن اشاره خواهد شد. همچنین فرض شد: $\gamma = 0.98$ و بردار حالت نیز به فرم زیر انتخاب گردید:

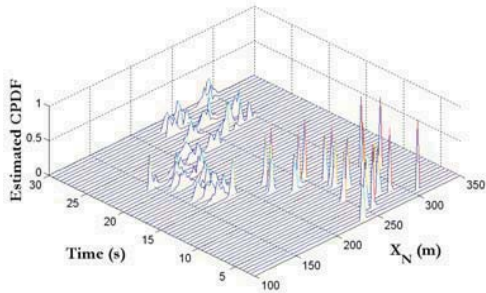
$$x = [X_N, X_E, X_D, V_N, V_E, V_D, roll, pitch, turn, b_n, b_e, b_d, b_w, b_p, b_r, x_r, \delta_r]$$

جهت ارزیابی و مقایسه، صرفاً حالت بحرانی در نظر گرفته می‌شود، بنابراین فرض می‌کنیم که در ثانیه ۱۴۸۰م یک ماهواره از افق دید گیرنده خارج می‌شود. در این حالت همانطور که در شکل (۴) دیده می‌شود، تخمین انجام شده توسط فیلتر کالمن، به سرعت و بفرم تقریباً نامایی، واگرا می‌شود و خطای تخمین در انتهای بازه بحرانی اصلاً قابل قبول نیست. این امر قابل پیش‌بینی بود و لذا معضل عدم رویت ماهواره‌ها یکی از مواردیست که عملکرد فیلتر کالمن در مقوله ترکیب اطلاعات سیستم ناوبری را به شدت مخدوش می‌کند. نقطه صفر محور زمان در این نمودار، منطبق بر شروع لحظه بحرانی و ثانیه ۱۴۸۰م است.

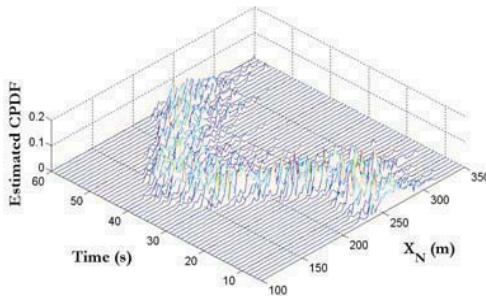
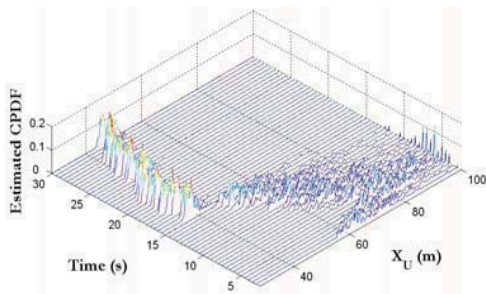


شکل ۷: خطای تخمین مکان با فرض رویت ۳ ماهواره توسط فیلتر کالمن توسعه یافته

این امر درحالیست که طبق شکل (۵) خطای روش پیشنهادی بسیار کمتر از فیلتر کالمن توسعه یافته است و می‌تواند جهت ناوبری در

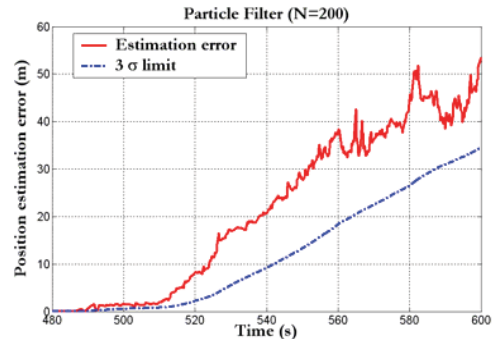


شکل ۹: تخمین تابع چگالی احتمال X_U و X_N در ۶۰ ثانیه اول حرکت به ازای $N=100$

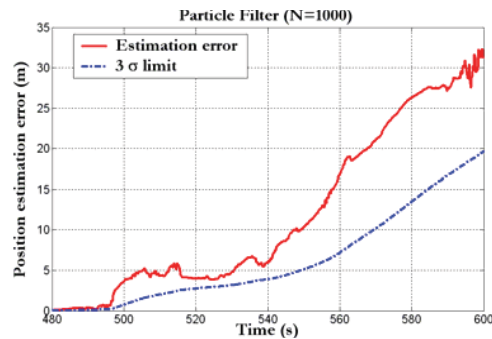


شکل ۱۰: تخمین تابع چگالی احتمال X_U و X_N در ۶۰ ثانیه اول حرکت به ازای $N=1000$

بالاخره آزمایش قطع ارتباط ماهواره‌ای برای ۴ مقدار متفاوت N انجام شد. مقادیر عددی متوسط و انحراف استاندارد خطای تخمین مکان در بازه بحرانی ذکر شده در جداول (۳) و (۴) آمده است. اختلاف عملکرد دو فیلتر با مقایسه اعداد، به خوبی قابل برداشت است. توجه شود که در عمل ممکن است در خیلی موارد حالت بحرانی پیش نیاید و تعداد ماهواره‌های تحت رویت، کمتر از مقدار بحرانی ۴ نشود و مثلاً ۸ یا ۹ ماهواره قابل رویت باشد. لیکن جهت اطمینان از حفظ عملکرد سیستم ناوبری شهری، در نظر گرفتن چنین شرایط خاصی الزامی است.

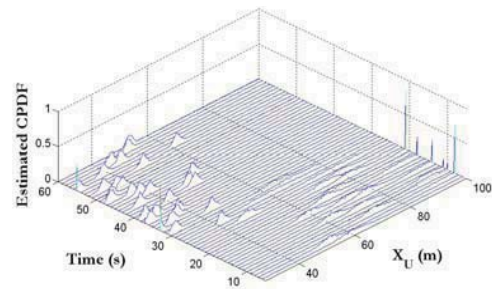


شکل ۷: خطای تخمین مکان با فرض رویت ۳ ماهواره توسط فیلتر ذره‌ای به ازای $N=200$



شکل ۸: خطای تخمین مکان با فرض رویت ۳ ماهواره توسط فیلتر ذره‌ای به ازای $N=1000$

در اشکال (۹) و (۱۰) نیز فرم زمانی تخمین تابع چگالی احتمال شرطی X_U و X_N به عنوان دو عضو از عناصر بردار حالت در ۶۰ ثانیه اول حرکت به ازای دو حالت اخیر رسم شده است که افزایش وضوح فیلتر در حالت دوم کاملاً مشهود می‌باشد.



جدول ۵: مقایسه حجم محاسبات مورد نیاز فیلتر ذره‌ای بر حسب شاخص فیلتر کالمن

توسعه یافته در حالت‌های بررسی شده

تعداد ذرات	۱۰۰	۲۰۰	۵۰۰	۱۰۰۰
تعداد فلاپس‌ها	۴/۹۳	۱۰/۲۲	۲۶/۷۸	۵۵/۰۴

۵- نتیجه‌گیری

با توجه به شبیه‌سازی‌های انجام شده، ملاحظه شد که در حالتی که تعداد ماهواره‌های قابل رویت کمتر از حداقل تعداد مورد نیاز برای مکان‌یابی ۳ بعدی می‌باشد، استفاده از فیلتر ذره‌ای برای تخمین حالت در سیستم ناوبری عملکرد مطلوبی دارد و خطا حداقل برای مدت مناسبی، محدود باقی می‌ماند درحالی‌که فیلتر کالمن سریعاً واگرا می‌شود. البته در حالت عادی و مشاهده همان حداقل ۴ ماهواره، می‌توان از فیلتر کالمن توسعه یافته استفاده کرد، ولی بلافاصله در زمانی که تعداد ماهواره‌های قابل رویت از حد مجاز کمتر شد، از الگوریتم پیشنهادی استفاده کرد.

البته همانطور که ذکر شده، این روش به توان محاسباتی بیشتری احتیاج دارد و زیاد بودن تعداد حالت‌ها در برخی موارد، امکان پردازش همزمان را در کاربردهای عملی سلب می‌کند که این مشکل با افزایش سرعت پردازشگرهای الکترونیکی تقریباً حل شده فرض می‌شود.

از طرفی این روش در مواجهه با سیستم‌هایی با درجه غیرخطی بالا، مستقیماً و بدون خطی‌سازی و صرفاً بر اساس مدل واقعی سیستم، بهینه‌سازی را انجام می‌دهد. در نتیجه می‌توان انتظار عملکرد کاملاً مطلوب‌تری نسبت به فیلتر کالمن توسعه یافته را متصور بود. این نکته به این دلیل ذکر شد که در دینامیک‌های شدید پروازی، مدل سیستم تلفیقی GPS/SDINS به شدت پیچیده‌تر از آنچه در این شبیه‌سازی مورد استفاده قرار گرفت، بوده و خطی‌سازی آن کار نسبتاً دشواری می‌باشد. پس در عمل می‌توان به ارزش روش بیان شده و برتری‌های آن بخصوص در موارد بحرانی نسبت به روش متداولی چون فیلتر کالمن توسعه یافته پی‌برد.

مراجع

- [1] M. Hilberg and T. Jacob, "High accuracy navigation and landing system using GPS/IMU system integration", *IEEE AES Systems Magazine*, pp. 11-17, July 1998.
- [2] J. F. Wagner and T. Wieneke Integrating, "Integrating satellite and inertial navigation-conventional and new fusion approaches", Elsevier

جدول ۳: خطای تخمین مکان با فرض رویت ۳ ماهواره (متر)

خطای متوسط	خطای انحراف استاندارد خطا	
۲۶/۲۷	۲۲/۸۴	N=100
۲۱/۷۰	۱۶/۹۴	N=200
۱۲/۹۱	۱۲/۷۳	N=500
۱۱/۸۹	۱۰/۱۱	N=1000
۵۹/۱۹۵	۴۸/۶۴۴	فیلتر کالمن توسعه یافته

جدول ۴: خطای تخمین مکان با فرض رویت ۲ ماهواره (متر)

خطای متوسط	خطای انحراف استاندارد خطا	
۱۴۶/۶۷	۱۳۷/۱۷	N=100
۹۸/۳۱	۹۰/۱۸	N=200
۵۰/۵۱	۴۰/۰۸	N=500
۴۲/۱۲	۳۶/۵۹	N=1000
۳۷۱/۶۴	۳۲۴/۳۵	فیلتر کالمن توسعه یافته

بدیهی می‌باشد که در حالت عادی و غیربحرانی نیز میزان صحت فیلتر ذره‌ای به مراتب از فیلتر کالمن بیشتر است (البته به شرط انتخاب تعداد ذرات کافی)، که در اینجا این امر مورد بحث ما نیست [۱۶]. اما هزینه این افزایش صحت در عملکرد سیستم ناوبری، زیاد شدن بار محاسباتی مساله است. اجرای کل برنامه تخمین فیلتر ذره‌ای در طول زمان ۱۱ دقیقه‌ای ناوبری بر روی یک PC با پردازنده AMD Athlon XP 1800+ و حافظه 512 MB (DDR 400 MHz) و به ازای $N=500$ ، حدود ۸۹/۹۷ ثانیه بطول انجامید. همین آزمایش برای فیلتر کالمن حدود ۳/۴۵۳ ثانیه طول کشید.

مقایسه حالت‌های بررسی شده طبق معیار فلاپس در برنامه MATLAB که ارتباط بسیار نزدیکی با توان پردازشی مورد نیاز الگوریتم دارد، در جدول (۵) نشان داده شده است. در این حالت تعداد فلاپس‌های مورد استفاده توسط فیلتر کالمن توسعه یافته به عنوان معیار، برابر واحد فرض شده است.

- Electronic Systems, vol 33, no. 3, pp. 835-850, July 1997.
- [11] S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking", IEEE Transactions on Signal Processing, vol 50, no 2, pp. 174-188, February 2002.
- [12] C. Hue, J. P. Le Cadre and P. Pérez, "Sequential monte carlo methods for multiple target tracking and data fusion", IEEE Transactions on Signal Processing, vol. 50, no. 2, pp. 309-325, February 2002.
- [13] J. H. Kotecha and P. M. Djuric, "Gaussian particle filtering", IEEE Transactions on Signal Processing, vol 51, no 10, pp. 2592- 2601, October 2003.
- [14] B. Azimi Sadjadi and P. S. Krishnaprasad, "Approximate nonlinear filtering and its application in navigation", Ph.D. Dissertation, Department of Electrical Engineering, Maryland University, College Park, 2001.
- [15] علی اسدیان، بهزاد مشیری و علی خاکی صدیق، "بهبودسازی غیرخطی با کمک فیلترهای ذره‌ای در سیستم تلفیقی GPS/INS در شرایط بحرانی"، سیزدهمین کنفرانس مهندسی برق ایران، مجموعه مقالات کنترل و مهندسی پزشکی، ص. ۹۳ الی ۹۹، ۲۰ الی ۲۲ اردیبهشت، زنجان، ایران.
- [16] علی اسدیان، "ارزیابی روش‌های کلاسیک و هوشمند ترکیب اطلاعات در سیستم تلفیقی GPS/INS"، پایان‌نامه کارشناسی ارشد مهندسی برق-کنترل، دانشکده برق و کامپیوتر دانشگاه تهران، ۱۳۸۴.
- [17] <http://www.gpssoftnav.com>
- Journal of Control Engineering Practice 11, pp. 543-550, 2003.
- [3] J. A. Farrel, and M. Barth, "The global positioning system and inertial navigation", McGraw-Hill, 1999.
- [4] X. He, Y. Chen and H.B.Iz, "A reduced-order model for integrated GPS/INS", IEEE AES Systems Magazine, pp. 40-45, March 1998.
- [5] S. Nasser, "Improving the inertial navigation system (INS) error model for INS and INS/DGPS applications", Ph.D. Dissertation, Department of Geomatics Engineering, Calgary University, Alberta, CA, November 2004.
- [6] G. Welsh and G. Bishop, "An introduction to the Kalman filter", SIGGRAPH 2001, Course 8, Aug. 2001.
- [7] F. A. Faruqi and K. J. Turner, "Extended Kalman filter synthesis for integrated global positioning/inertial navigation systems", Elsevier Journal of Applied Mathematics and Computation 115, pp. 213-227, 2000.
- [8] D. M. Mayhew, "Multi-rate Sensor Fusion for GPS Navigation Using Kalman Filtering", M.S. Thesis in Electrical Engineering, Virginia Polytechnic Institute and State University, May 1999.
- [9] C. Hide, T. Moore and M. Smith, "Adaptive Kalman filtering for low-cost INS/GPS", Journal of Navigation, pp. 143-152, Cambridge University Press, 2003.
- [10] H. Carvalho, P. Del Moral, A. Monin and G. Salut, "Optimal nonlinear filtering in GPS/INS integration", IEEE Transactions on Aerospace

طراحی مطمئن کنترل کننده H_∞ برای سیستم های خطی تأخیر دار با وجود

نامعینی در تمام ماتریس های سیستم

علی ابوطالبی^۱، سعید حسین نیا^۲، فرید شیخ الاسلام^۳

^۱ کارشناسی ارشد، دانشگاه صنعتی اصفهان، abootalebi_a@alumni.iut.ac.ir

^۲ استادیار، دانشگاه صنعتی اصفهان، hoseinia@cc.iut.ac.ir

^۳ دانشیار، دانشگاه صنعتی اصفهان، sheikh@cc.iut.ac.ir

چکیده: موضوع این مقاله طراحی مطمئن کنترل کننده مقاوم H_∞ برای سیستم های خطی نامعین دارای تأخیر زمانی با وجود خرابی در محرک های کنترل می باشد. فرض می شود نامعینی های پارامتری با نرم محدود در تمام ماتریس های معادلات حالت حضور داشته باشند و خرابی در یک زیر مجموعه معین از محرکها اتفاق می افتد. برای حل این مسئله یک معادله جبری ریکاتی توسعه یافته پیشنهاد می شود که در نهایت با حل آن و اعمال کنترل کننده بدست آمده توسط آن به سیستم، پایداری و رفتار H_∞ هم در حالت عادی و هم در مواقع وقوع خطا در محرک های کنترل تضمین خواهد شد.

واژه های کلیدی: کنترل مطمئن، کنترل H_∞ خطی، سیستم های نامعین، تأخیر زمانی.

Abstract: This paper deals with the reliable H-infinity controller design for linear uncertain time-delay systems in the presence of actuator failures. It is assumed that the norm bounded uncertainties are present in all system matrices and the actuator failures occur among a prespecified subset of actuators. A modified algebraic Riccati equation is derived whose solution provides the desired controller. The proposed controller guarantees robust stability and H-infinity performance not only when all control components are operational, but also in the case of actuator failures.

Keywords: Reliable Control, Linear H-infinity Control, Uncertain Systems, Time-Delay

تضمین می کند ولیکن باید توجه داشت که در طراحی سیستم های کنترل معمولاً فرض بر این است که تمام سنسورها و محرک ها در شرایط کاری مناسبی قرار دارند در صورتی که در کاربردهای عملی معمولاً خطا در المانهای کنترل اتفاق می افتد و در هنگام وقوع خطا ممکن است این روش کارایی خود را از دست داده و موجب اختلال در رفتار سیستم و یا حتی ناپایداری سیستم گردد. به این علت تلاشهایی نیز بعمل آمده که بتوان سیستم کنترلی طراحی نمود که بتواند خطاهای رخ داده در بخشی از محرکها و سنسورهای کنترل را تحمل کند. این مسئله به طراحی مطمئن سیستم های کنترل معروف است.

در گذشته تلاشهای بسیاری در این زمینه صورت گرفته و سعی شده با روشهای مختلف طراحی مطمئنی برای سیستم های کنترل انجام گیرد. در این رابطه در [۱۴] روشی ارائه شده که در آن از یک کنترل کننده ثانویه استفاده شده و مجموعه این دو کنترل کننده سیستم را پایدار می کند. همچنین در [۱۵] روشی مطرح شده که در آن از یک کنترل کننده وقعی که به صورت موازی با یک کنترل کننده ثابت کار می کند استفاده شده و

۱- مقدمه

در یک دهه گذشته روش H_∞ به عنوان یک روش مؤثر در حل مسائلی مانند حذف اثر اغتشاش، ردیابی و طراحی مقاوم مورد استفاده قرار گرفته که مروری بر آن در [۱] ارائه شده است. به علاوه به منظور بهبود رفتار مقاوم در برابر نامعینی های پارامتری مسئله کنترل H_∞ مقاوم نیز مورد مطالعه قرار گرفته است [۲-۸]. از طرفی به علت اینکه در رفتار دینامیکی اکثر سیستم های فیزیکی به طور ذاتی تأخیر زمانی وجود دارد، روش مذکور برای سیستم های با تأخیر زمانی توسعه داده شده و مقالات متعددی در این رابطه به چاپ رسیده است که به عنوان مثال میتوان به [۹-۱۲] اشاره کرد.

بعلاوه این روش در مورد سیستم هایی که هم دارای نامعینی و هم دارای تأخیر زمانی باشند بررسی شده است [۱۳]. اگر چه روش طراحی کنترل کننده مقاوم H_∞ در مورد سیستم های ذکر شده در حالت کار عادی روش مناسبی است و پایداری و رفتار مناسب این سیستم ها را

نشان دادن یک ماتریس افزوده شده توسط ماتریس های A, B استفاده می گردد.

۲- توصیف سیستم

یک سیستم خطی نامعین دارای تأخیر زمانی را بصورت زیر در نظر بگیرد:

$$\begin{aligned} \dot{x}(t) &= (A + \Delta A)x(t) + (A_d + \Delta A_d)x(t-h) \\ &+ (B + \Delta B)u(t) + Dw(t) \\ z(t) &= \begin{bmatrix} (C + \Delta C)x(t) \\ u(t) \end{bmatrix} \\ x(t) &= \phi(t), t \in [-h, 0] \end{aligned} \quad (1)$$

که $x(t) \in R^n$ حالت سیستم، $u(t) \in R^m$ ورودی کنترل، $w(t) \in R^p$ ورودی اغتشاش و $z(t)$ خروجی کنترل شده می باشد. A, A_d, B, C, D ماتریس های حقیقی ثابت توصیف کننده سیستم نامی با بعد مناسب می باشند. همچنین h معرف تأخیر حالت و $\phi(t)$ یک بردار پیوسته معرف مقدار اولیه می باشد، همچنین $\Delta A, \Delta A_d, \Delta B, \Delta C$ نامعینی های پارامتری با نرم محدود می باشند که به صورت زیر توصیف می شوند:

$$\begin{bmatrix} \Delta A & \Delta A_d & \Delta B \\ \Delta C & 0 & 0 \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \end{bmatrix} F \begin{bmatrix} E_1 & E_2 & E_3 \end{bmatrix} \quad (2)$$

که

$$H_1 \in R^{n \times r}, H_2 \in R^{l \times r}, E_1 \in R^{q \times n}, E_2 \in R^{q \times n}, E_3 \in R^{q \times m}$$

ماتریس های ثابت معرف ساختار نامعینی ها بوده و همچنین $F \in R^{r \times q}$ یک ماتریس نامعین است که شرایط زیر را برآورده می سازد:

$$FF^T \leq \bar{F}_1, F^T F \leq \bar{F}_2 \quad (3)$$

که در این رابطه \bar{F}_1 و \bar{F}_2 بصورت حد بالای نامعینی تعریف میشوند $\Delta A, \Delta A_d, \Delta B, \Delta C$ قابل قبول خوانده می شوند اگر (۲) و (۳) برآورده شوند.

در اینجا هدف طراحی یک کنترل کننده خطی فیدبک حالت به فرم

$$u(t) = Kx(t)$$

می باشد به گونه ای که با اعمال این کنترل کننده به سیستم (۱) پایداری و رفتار مناسب سیستم از نظر محدوده نرم H_∞ هم در حالت کارکرد عادی و هم به هنگام خرابی بعضی از محرک های کنترل تضمین گردد. روند کلی در این باره تفکیک محرک های کنترل به دو گروه می باشد [۲۵] که گروه اول خطا را تحمل کرده و گروه دوم تحمل خطا را ندارند. بدین

در مدهای خطای مشخصی پایداری تضمین میشود. روشی دیگر انتخاب از میان بهره های فیدبک حالت پایداری ساز است بگونه ای که علی رغم وجود سنسورهای خراب پایداری را تضمین کند. لازم به ذکر است که در این روشها و چند روش ارائه شده دیگر [۱۶-۱۹] به موضوع تضمین رفتار مناسب سیستم پرداخته نشده است. در ادامه این روند این موضوع با بکارگیری روش کنترل H_∞ در طراحی مطمئن سیستم های کنترل مورد توجه قرار گرفت.

در این رابطه در [۲۰] یک روش طراحی مطمئن برای سیستم های متمرکز و غیر متمرکز مورد بحث قرار گرفت که در آن به مسئله پایداری و تضعیف اغتشاش توأم پرداخته شده ولی برای سیستم نامعینی در نظر گرفته نشده است. علاوه بر این روش برای سیستم های نامعین مورد بررسی قرار گرفته [۲۱ و ۲۲] و در ادامه برای سیستم های دارای تأخیر زمانی بدون نامعینی نیز توسعه داده شده است [۲۳]. آخرین نتیجه در این زمینه مربوط به سیستم هایی است که دارای تأخیر زمانی هستند اما نامعینی در آنها فقط در یکی از ضرایب ماتریسی ظاهر می شود [۲۴] و تاکنون این روش برای سیستم هایی که هم دارای تأخیر زمانی و هم دارای نامعینی در تمام ماتریسها باشند مطرح نشده است.

موضوع این مقاله طراحی مطمئن سیستم کنترل با روش H_∞ برای سیستم های خطی است که هم دارای تأخیر زمانی و هم دارای نامعینی در کلیه ماتریس های معادله حالت باشند. هدف نهایی طرح این مسئله طراحی یک کنترل کننده فیدبک حالت می باشد که پایداری و رفتار مناسب سیستم را علی رغم حضور نامعینی در تمام ماتریس های معادله حالت و تأخیر زمانی و همچنین وجود خطا در محرک های کنترل تضمین کند. در این مسئله فرض می شود نامعینی های پارامتری دارای نرم محدودی باشند و خطا در یک زیر مجموعه معین از محرک های کنترل رخ دهد. باید خاطر نشان کرد که روش پیشنهادی در این مقاله مبتنی بر یک معادله جبری ریگاتی توسعه یافته است.

مقاله حاضر مشتمل بر پنج بخش است که پس از مقدمه، بخش دوم به معرفی مدل سیستم مورد بحث می پردازد. در بخش سوم به طراحی مطمئن کنترل کننده در حضور خطا در گروهی از محرک های کنترل پرداخته می شود. بخش چهارم به نتایج شبیه سازی برای یک مثال عددی اختصاص دارد و در پایان در بخش پنجم به نتیجه گیری می پردازیم.

علامت بکار رفته در این مقاله کاملاً استاندارد هستند. بطور خاص، بالانویس "T" نشان دهنده ترانپوز است و نامعادله $X > Y$ (یا $X \geq Y$) مثبت معین (یا مثبت نیم معین) بودن $X - Y$ را نشان می دهد که X, Y دو ماتریس متقارن هستند. از علامت $\begin{bmatrix} A & B \end{bmatrix}$ برای

$$\Psi = [C^T H_2 (I - H_2^T H_2)^{-\frac{1}{2}} - E_1^T F^T (I - H_2^T H_2)^{\frac{1}{2}}]$$

که با توجه به نامعادله

$$\Psi \Psi^T \geq 0$$

خواهیم داشت

$$\begin{aligned} 0 &\leq [C^T H_2 (I - H_2^T H_2)^{-\frac{1}{2}} - E_1^T F^T (I - H_2^T H_2)^{\frac{1}{2}}] \\ &\times [C^T H_2 (I - H_2^T H_2)^{-\frac{1}{2}} - E_1^T F^T (I - H_2^T H_2)^{\frac{1}{2}}]^T \\ &= C^T H_2 (I - H_2^T H_2)^{-1} H_2^T C \\ &+ E_1^T F^T (I - H_2^T H_2) F E_1 - C^T H_2 F E_1 - E_1^T F^T H_2^T C \\ &= C^T H_2 (I - H_2^T H_2)^{-1} H_2^T C + E_1^T F^T F E_1 \\ &- E_1^T F^T H_2^T H_2 F E_1 - C^T H_2 F E_1 - E_1^T F^T H_2^T C \\ &= C^T H_2 (I - H_2^T H_2)^{-1} H_2^T C + E_1^T F^T F E_1 \\ &- (\Delta C)^T (\Delta C) - C^T (\Delta C) - (\Delta C)^T C \end{aligned}$$

بنابراین با توجه به روابط فوق و شرط $F^T F \leq \bar{F}_2$ داریم

$$\begin{aligned} &(\Delta C)^T (\Delta C) + C^T (\Delta C) + (\Delta C)^T C \\ &\leq C^T H_2 (I - H_2^T H_2)^{-1} H_2^T C + E_1^T \bar{F}_2 E_1 \end{aligned}$$

که در نهایت با اضافه کردن $C^T C$ به طرفین این رابطه به نامعادله (۸) خواهیم رسید.

لم ۳. اگر ماتریس مثبت معین Q وجود داشته باشد بگونه ای که

$$\begin{aligned} &I - E_2 Q^{-1} E_2^T \\ &(A_d + \Delta A_d) Q^{-1} (A_d + \Delta A_d)^T \leq A_d Q^{-1} A_d^T \quad (9) \\ &+ A_d Q^{-1} E_2^T (I - E_2 Q^{-1} E_2^T)^{-1} E_2 Q^{-1} A_d^T + H_1 \bar{F}_1 H_1^T \quad (10) \end{aligned}$$

اثبات: مشابه لم ۲ با این تفاوت که در این حالت

$$\begin{aligned} \Psi &= [A_d Q^{-1} E_2^T (I - E_2 Q^{-1} E_2^T)^{-\frac{1}{2}} \\ &- H_1 F (I - E_2 Q^{-1} E_2^T)^{\frac{1}{2}}] \end{aligned}$$

و همچنین شرط $FF^T \leq \bar{F}_1$ در نظر گرفته می شود.

لم ۴. با فرض $I - E_3 E_3^T > 0$ نامعادله زیر برقرار است.

$$(\Delta B)(\Delta B)^T \leq H_1 \bar{F}_1 H_1^T \quad (10)$$

اثبات: مشابه لم ۳ و با فرض

$$\Psi = [H_1 F (I - E_3 E_3^T)^{\frac{1}{2}}]$$

لم ۵. فرض کنید سیستم (۱) شرایط (۲) و (۳) را برآورده می سازد. در این صورت برای هر $\varepsilon_1 > 0, \varepsilon_2 > 0$ که ε_1 و ε_2 اسکالر باشند نامعادلات زیر را داریم:

$$-B_{\bar{\omega}} (\Delta B)^T - (\Delta B) B_{\bar{\omega}}^T \leq \varepsilon_1^{-1} B_{\bar{\omega}} B_{\bar{\omega}}^T + \varepsilon_1 H_1 \bar{F}_1 H_1^T \quad (11)$$

ترتیب می توان دو مجموعه متناظر با این دو گروه را بصورت زیر تعریف کرد:

$$\bar{\Omega} \subseteq \{1, 2, \dots, m\} - \Omega, \quad \Omega \subseteq \{1, 2, \dots, m\}$$

که Ω معرف محرکهای مشکوک به خرابی و $\bar{\Omega}$ نمایانگر محرکهای است که خرابی در آنها رخ نمی دهد. با توجه به این دسته بندی می توان از تجزیه زیر استفاده کرد:

$$B = B_{\Omega} + B_{\bar{\Omega}} \quad (4)$$

که B_{Ω} و $B_{\bar{\Omega}}$ از صفر کردن ستونهای متناظر با $\bar{\Omega}$ و Ω در ماتریس B بدست می آیند. همچنین اگر $\omega \subseteq \Omega$ متناظر با آن دسته از محرکهای کنترل باشد که در آنها واقعاً خطا رخ داده است می توان بطور مشابه تجزیه زیر را تعریف کرد:

$$B = B_{\omega} + B_{\bar{\omega}} \quad (5)$$

که B_{ω} و $B_{\bar{\omega}}$ دارای تعریفی مشابه با B_{Ω} و $B_{\bar{\Omega}}$ می باشند.

۳- طراحی کنترل کننده

برای طراحی کنترل کننده و اثبات قضیه اصلی از لم های زیر استفاده خواهد شد.

لم ۱. [۱۲] اگر γ یک ثابت مثبت و Q یک ماتریس مثبت معین بوده و یک ماتریس مثبت معین P وجود داشته باشد بطوری که نامعادله زیر برآورده شود:

$$P - A Q^{-1} A^T - H_1 \bar{F}_1 H_1^T > 0 \quad (6)$$

آنگاه سیستم

$$\dot{x}(t) = \hat{A}x(t) + A_d x(t-h) + \hat{D}\hat{w}(t) \quad (7)$$

$$z(t) = \hat{C}x(t)$$

پایدار مجانبی بوده و برای تابع تبدیل

$$H_{zw}(s) = \hat{C}(sI - \hat{A} - A_d e^{-sh})^{-1} \hat{D}$$

از ورودی اغتشاش $\hat{w}(t)$ به خروجی تحت کنترل $z(t)$ نامعادله

$$\|H_{zw}(s)\|_{\infty} \leq \gamma$$

برای هر $h > 0$ برقرار است.

تبصره ۱: چون لم (۱) مستقل از مقدار تأخیر اثبات می گردد از این لم می توان برای حالتی که مقدار تأخیر زمانی نامعلوم است نیز استفاده کرد.

لم ۲. با فرض اینکه $I - H_2^T H_2$ مثبت معین باشد نامعادله زیر برقرار است:

$$(C + \Delta C)^T (C + \Delta C) \leq C^T C + \quad (8)$$

اثبات: ماتریس Ψ را بصورت زیر تعریف می کنیم

اثبات: برای هر دو ماتریس X و Y با بعد مناسب و برای هر $\varepsilon > 0$ $-B_\omega(\Delta B)^T - (\Delta B)B_\omega^T \leq \varepsilon_2^{-1}B_\omega B_\omega^T + \varepsilon_2 H_1 \bar{F}_1 H_1^T$ $\omega \subseteq \Omega$ نرم H_∞ تابع انتقال حلقه بسته کوچکتر یا مساوی γ خواهد شد.

اثبات: با قرار دادن قانون کنترل (۱۴) در (۱) داریم

$$\begin{aligned} \dot{x}(t) &= (A + \Delta A + (B + \Delta B)K)x(t) \\ &+ (A_d + \Delta A_d)x(t-h) + Dw(t) \\ z(t) &= \begin{bmatrix} (C + \Delta C)x(t) \\ u(t) \end{bmatrix} \end{aligned}$$

$$x(t) = \phi(t), \quad t \in [-h, 0]$$

که با استفاده از (۵) خواهیم داشت

$$\begin{aligned} (B + \Delta B)K &= -(B_\omega + B_{\bar{\omega}} + \Delta B)(B_\omega + B_{\bar{\omega}})^T P \\ &= -B_\omega B_\omega^T P - B_{\bar{\omega}} B_{\bar{\omega}}^T P - (\Delta B)B_\omega^T P - (\Delta B)B_{\bar{\omega}}^T P \end{aligned}$$

باید توجه داشت که ورودی کنترل $u(t)$ فقط مربوط به محرکهای سالم می باشد و خروجی های محرکهای خراب به منزله اغتشاش اضافی محسوب می شوند. از این رو داریم:

$$\begin{aligned} \dot{x}(t) &= A_c x(t) + A_1 x(t-h) + D_c w_c(t) \quad (15) \\ z(t) &= C_c x(t) \end{aligned}$$

که

$$\begin{aligned} A_c &= A + \Delta A - B_\omega B_\omega^T P - (\Delta B)B_\omega^T P - (\Delta B)B_{\bar{\omega}}^T P \\ A_1 &= A_d + \Delta A_d \quad C_c = \begin{bmatrix} C + \Delta C \\ -B^T P \end{bmatrix} \quad D_c = [D \quad B_\omega] \\ w_c(t) &= \begin{bmatrix} w(t) \\ v(t) \end{bmatrix} \quad v(t) = -B_\omega^T P x(t) \end{aligned}$$

اکنون بسادگی می توان از لم های گفته شده نامعادلات زیر را اثبات کرد:

$$\begin{aligned} A_c^T P + P A_c &\leq A^T P + P A + P[\varepsilon_3 H_1 H_1^T - 2B_\omega B_\omega^T P \\ &+ \varepsilon_1^{-1} B_\omega B_\omega^T + \varepsilon_1 H_1 \bar{F}_1 H_1^T + \varepsilon_2^{-1} B_\omega B_\omega^T + \varepsilon_2 H_1 \bar{F}_1 H_1^T] P \\ &+ \varepsilon_3^{-1} E_1^T \bar{F}_2 E_1 \quad (16) \\ P A_1 Q^{-1} A_1^T P &= P(A_d + \Delta A_d)Q^{-1}(A_d + \Delta A_d)^T P \\ &\leq P[A_d Q^{-1} A_d^T + A_d Q^{-1} E_2^T (I - E_2 Q^{-1} E_2^T)^{-1} E_2 Q^{-1} A_d^T \\ &+ H_1 \bar{F}_1 H_1^T] P \quad (17) \end{aligned}$$

و همچنین

$$\begin{aligned} C_c^T C_c &= (C + \Delta C)^T (C + \Delta C) + P B B^T P \quad (18) \\ &\leq C^T C + C^T H_2 (I - H_2^T H_2)^{-1} H_2 C + E_1^T \bar{F}_2 E_1 \\ &+ P B_\omega B_\omega^T P + P B_{\bar{\omega}} B_{\bar{\omega}}^T P \\ \gamma^{-2} P D_c D_c^T P &= \gamma^{-2} P D D^T P + \gamma^{-2} P B_\omega B_\omega^T P \quad (19) \end{aligned}$$

و از طرفی با توجه به

$$\begin{aligned} B_\Omega B_\Omega^T &= B_\omega B_\omega^T + B_{\Omega-\omega} B_{\Omega-\omega}^T \\ B_{\bar{\Omega}} B_{\bar{\Omega}}^T &= B_{\bar{\omega}} B_{\bar{\omega}}^T - B_{\Omega-\omega} B_{\Omega-\omega}^T \end{aligned}$$

$$XY^T + YX^T \leq \frac{1}{\varepsilon} XX^T + \varepsilon YY^T$$

اثبات این لم با توجه به نامعادله بالا و لم (۴) امکان پذیر است.

اکنون با این فرض که خروجی محرکهای خراب شده متعلق به فضای $L_2[0, \infty)$ باشند اثر این محرکها را به صورت یک ورودی اغتشاش بنام $v(t)$ بغیر از $w(t)$ در نظر می گیریم. بنابراین با تعریف $w_c = [w^T \ v^T]^T$ می توان سیستم تحت کنترل را به فرم زیر نشان داد.

$$\begin{aligned} \dot{x}(t) &= (A + \Delta A)x(t) + (A_d + \Delta A_d)x(t-h) \quad (12) \\ &+ (B_\omega + \Delta B)u(t) + [D \quad B_\omega]w_c(t) \end{aligned}$$

$$z(t) = \begin{bmatrix} (C + \Delta C)x(t) \\ u(t) \end{bmatrix}$$

$$x(t) = \phi(t), \quad t \in [-h, 0]$$

در ادامه ابتدا به معرفی معادله ریکاتی توسعه یافته زیر می پردازیم

$$\begin{aligned} A^T P + P A + P \Sigma P + (1 + \varepsilon_3^{-1}) E_1^T \bar{F}_2 E_1 + C^T C \quad (13) \\ + C^T H_2 (I - H_2^T H_2)^{-1} H_2^T C + Q + Q_1 = 0 \end{aligned}$$

که

$$\begin{aligned} \Sigma &:= \varepsilon_3 H_1 H_1^T + (-1 + \varepsilon_1^{-1}) B_{\bar{\Omega}} B_{\bar{\Omega}}^T \\ &+ (1 + \varepsilon_1 + \varepsilon_2) H_1 \bar{F}_1 H_1^T + A_d Q^{-1} A_d^T \\ &+ A_d Q^{-1} E_2^T (I - E_2 Q^{-1} E_2^T)^{-1} E_2 Q^{-1} A_d^T \\ &+ (1 + \gamma^{-2} + \varepsilon_2^{-1}) B_\Omega B_\Omega^T + \gamma^{-2} D D^T \end{aligned}$$

و $\varepsilon_1, \varepsilon_2, \varepsilon_3$ اسکالرهایی مثبت و Q_1 و Q ماتریس های مثبت معین می باشند.

قضیه. اگر $\gamma > 0$ یک ثابت حقیقی باشد و همچنین ماتریس های H_2 و E_3 نامعادلات

$$H_2^T H_2 < I, E_3 E_3^T < I$$

را برآورده کرده و اسکالرهایی مثبت $\varepsilon_2 > 0$ ، $\varepsilon_1 > I$ ، و $\varepsilon_3 > 0$

ماتریس های مثبت معین $Q > 0, Q_1 > 0$ وجود داشته باشند بطوریکه

$$E_2 Q^{-1} E_2^T < I$$

و معادله ریکاتی (۱۳) دارای یک حل مثبت معین $P > 0$ باشد. آنگاه

قانون کنترل فیدبک حالت

$$u(t) = Kx(t), \quad K = -B^T P \quad (14)$$

سیستم نامعین با تأخیر زمانی (۱) را برای هر $h > 0$ پایدار کرده و برای

تمام نامعینی های قابل قبول و همچنین خطاهای محرکهای متناظر با

می توان نوشت:

$$B_{\bar{\Omega}} B_{\bar{\Omega}}^T \leq B_{\bar{\omega}} B_{\bar{\omega}}^T \quad (20)$$

$$B_{\omega} B_{\omega}^T \leq B_{\Omega} B_{\Omega}^T$$

سرانجام از روابط بالا و معادله ریکاتی (۶) داریم

$$\begin{aligned} & A_c^T P + PA + PA_d Q^{-1} A_d^T P + Q + C_c^T C_c \\ & + \gamma^{-2} P D_c D_c^T P A^T P \leq A^T P + PA \\ & + P[\varepsilon_3 H_1 H_1^T - 2B_{\bar{\omega}} B_{\bar{\omega}}^T + \varepsilon_1^{-1} B_{\bar{\omega}} B_{\bar{\omega}}^T \\ & + \varepsilon_1 H_1 \bar{F}_1 H_1^T + \varepsilon_2^{-1} B_{\omega} B_{\omega}^T + \varepsilon_2 H_1 \bar{F}_1 H_1^T] P \\ & + \varepsilon_3^{-1} E_1^T \bar{F}_2 E_1 + P[A_d Q^{-1} A_d^T \\ & + A_d Q^{-1} E_2^T (I - E_2 Q^{-1} E_2^T)^{-1} E_2 Q^{-1} A_d^T \\ & + H_1 \bar{F}_1 H_1^T] P + Q + C^T C \\ & + C^T H_2 (I - H_2^T H_2)^{-1} H_2 C + E_1^T \bar{F}_2 E_1 \\ & + P B_{\bar{\omega}} B_{\bar{\omega}}^T P + P B_{\omega} B_{\omega}^T P + \gamma^{-2} P D D^T P \\ & + \gamma^{-2} P B_{\omega} B_{\omega}^T P \\ & = \\ & A^T P + PA \\ & + P[\varepsilon_3 H_1 H_1^T + (-1 + \varepsilon_1^{-1}) B_{\bar{\omega}} B_{\bar{\omega}}^T \\ & (1 + \varepsilon_1 + \varepsilon_2) H_1 \bar{F}_1 H_1^T + (1 + \gamma^{-2} + \varepsilon_2^{-1}) B_{\omega} B_{\omega}^T \\ & + A_d Q^{-1} A_d^T + A_d Q^{-1} E_2^T (I - E_2 Q^{-1} E_2^T)^{-1} E_2 Q^{-1} A_d^T \\ & + \gamma^{-2} D D^T] P + (1 + \varepsilon_3^{-1}) E_1^T \bar{F}_2 E_1 + Q + C^T C \\ & + C^T H_2 (I - H_2^T H_2)^{-1} H_2 C \\ & \leq A^T P + PA \\ & + P[\varepsilon_3 H_1 H_1^T + (-1 + \varepsilon_1^{-1}) B_{\bar{\Omega}} B_{\bar{\Omega}}^T \\ & (1 + \varepsilon_1 + \varepsilon_2) H_1 \bar{F}_1 H_1^T + (1 + \gamma^{-2} + \varepsilon_2^{-1}) B_{\Omega} B_{\Omega}^T \\ & + A_d Q^{-1} A_d^T + A_d Q^{-1} E_2^T (I - E_2 Q^{-1} E_2^T)^{-1} E_2 Q^{-1} A_d^T \\ & + \gamma^{-2} D D^T] P + (1 + \varepsilon_3^{-1}) E_1^T \bar{F}_2 E_1 + Q + C^T C \\ & + C^T H_2 (I - H_2^T H_2)^{-1} H_2 C \\ & = A^T P + PA + P \Sigma P + (1 + \varepsilon_3^{-1}) E_1^T \bar{F}_2 E_1 + C^T C \\ & + C^T H_2 (I - H_2^T H_2)^{-1} H_2^T C + Q + Q_1 = 0 \end{aligned}$$

که با توجه به لم ۱ اثبات قضیه کامل می شود.

تبصره ۲: بدیهی است که حل مسئله مورد بحث به وجود یک پاسخ مثبت معین برای معادله (۱۳) بستگی دارد. بسادگی می توان مشاهده کرد که اگر \bar{F}_1 متقارن انتخاب شود ماتریس Σ در (۱۳) متقارن و مثبت نیم معین خواهد بود. در این صورت، با فرض تقارن \bar{F}_2 معادله (۱۳) به شکل معادله جبری ریکاتی تعمیم یافته درمی آید که حل آن از طریق روشهای عددی موجود بدست می آید.

باید توجه داشت که \bar{F}_1 و \bar{F}_2 (حد بالای نامعینی ها) را همیشه می توان متقارن انتخاب نمود. همچنین در صورتی که پارامترهای $\varepsilon_1, \varepsilon_2, \varepsilon_3$ بنحوی انتخاب شوند که ماتریس Σ مثبت معین باشد، شرایط لازم و کافی برای وجود یک حل متقارن و مثبت معین برای (۱۳) و همچنین حل مربوطه در [۲۴] ارائه شده است.

۴- مثال عددی

در این قسمت برای مشخص کردن کاربرد این روش یک مثال عددی ارائه می شود. سیستم با تأخیر زمانی نامعین زیر را در نظر بگیرید.

$$A = \begin{bmatrix} 6 & 0.5 & -0.5 \\ -1 & 2 & 0.2 \\ 1 & -2 & -6 \end{bmatrix}, A_d = \begin{bmatrix} -0.05 & 0 & 0.02 \\ 0.1 & -0.06 & 0 \\ -0.01 & -0.02 & -0.5 \end{bmatrix}$$

$$B = \begin{bmatrix} 3 & 0.1 & 0 \\ 1 & 3 & 0.1 \\ 0.2 & 0 & 0.3 \end{bmatrix}, C = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 0.4 & 0.1 \end{bmatrix}$$

$$D = \begin{bmatrix} 0.1 & 0 \\ 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, \bar{F}_1 = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$$

$$\bar{F}_2 = \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.3 \end{bmatrix}, H_1 = \begin{bmatrix} 0.08 & 0 & 0 \\ 0 & 0.08 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, E_1 = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$$

$$F = \text{diag}\{0.1, 0.1, 0.2\}$$

$$\phi(t) = [5 \quad -5 \quad 1], t \in [-h, 0]$$

$$E_2 = \begin{bmatrix} 0.04 & 0 & 0 \\ 0 & 0.02 & 0 \\ 0 & 0 & 0.04 \end{bmatrix}, E_3 = \begin{bmatrix} 0.02 & 0 & 0 \\ 0 & 0.02 & 0 \\ 0 & 0 & 0.02 \end{bmatrix}$$

مقادیر ویژه A عبارت است از:

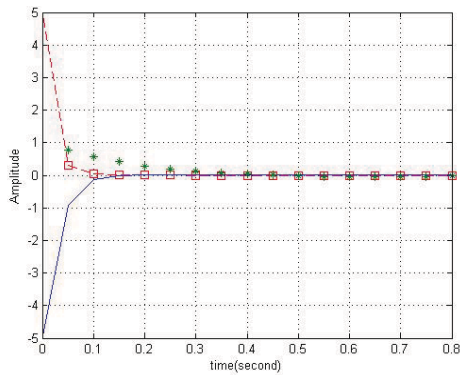
$$\{5.8080, 2.1095, -5.9175\}$$

و همانطور که مشاهده می شود A ناپایدار است. برای $\Omega = \{3\}$ ماتریس B را بصورت زیر تجزیه می کنیم.

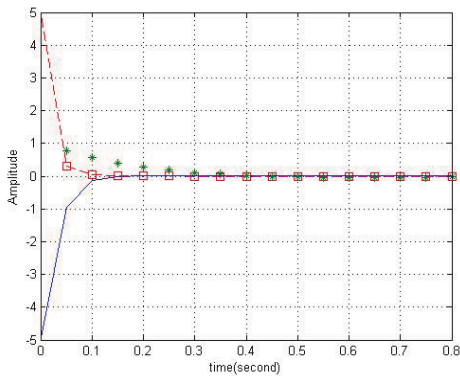
$$B_{\Omega} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0.1 \\ 0 & 0 & 0.3 \end{bmatrix}, B_{\bar{\Omega}} = \begin{bmatrix} 3 & 0.1 & 0 \\ 1 & 3 & 0 \\ 0.2 & 0 & 0 \end{bmatrix}$$

همچنین با انتخاب

$$Q = Q_1 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, \varepsilon_1 = \frac{3}{2}, \varepsilon_2 = 1, \varepsilon_3 = \frac{4}{3}$$

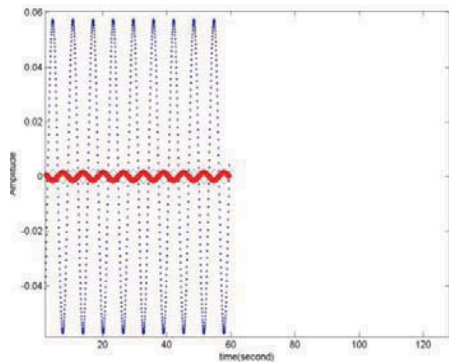


شکل ۱ پاسخ حالت برای حالتی که خط رخ ندهد و با فرض $\gamma = 0.1, h = 0.5$
 X_1 (مربع)، X_2 (خط) و X_3 (ستاره)



شکل ۲ پاسخ حالت برای حالتی که محرک سوم دچار خطا شود، و
 $\gamma = 0.1, h = 0.5$

X_1 (مربع)، X_2 (خط) و X_3 (ستاره)



شکل ۳ تضعیف اغتشاش برای حالتی که محرک سوم دچار خطا شود و

$\gamma = 0.1, h = 0.5$

X_1 (نقطه)، X_2 (خط) و X_3 (ستاره)

براحتی می توان مشاهده کرد که شرایط

$$H_2^T H_2 < I, E_2 Q^{-1} E_2^T < I, E_3 E_3^T < I$$

برقرار است. با انتخاب $\gamma = 0.1$ می توان ماتریس Σ و سپس جواب مثبت معین و متقارن معادله ریکاتی توسعه داده شده (۱۳) را محاسبه کرد که به صورت زیر بدست می آید.

$$\Sigma = \begin{bmatrix} -1.9608 & -0.1508 & -0.2965 \\ -0.1508 & -1.1625 & 2.9952 \\ -0.2965 & 2.9952 & 12.7123 \end{bmatrix}$$

$$P = \begin{bmatrix} 6.9736 & -0.9088 & -0.3070 \\ -0.9088 & 4.1836 & 0.1877 \\ -0.3070 & 0.1877 & 0.0524 \end{bmatrix}$$

بنابراین قانون کنترل فیدبک حالت بصورت زیر خواهد بود.

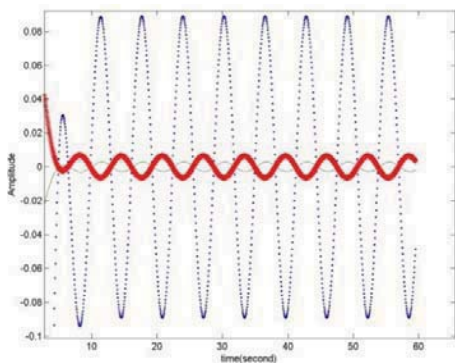
$$u(t) = Kx(t)$$

$$K = \begin{bmatrix} -19.9508 & -1.4949 & 0.7228 \\ 2.0289 & -12.4600 & -0.5323 \\ 0.1830 & -0.4747 & -0.0345 \end{bmatrix}$$

در ادامه با فرض اغتشاش سینوسی، پاسخ حالت سیستم را در دو حالت بدون خطا در محرکها و خرابی در سومین محرک رسم می نمایم تا از این طریق صحت نتایج بدست آمده را مورد بررسی قرار دهیم. به منظور بررسی تاثیر γ ، شبیه سازی ها برای دو مقدار $\gamma = 0.1$ و $\gamma = 3.1$ صورت می گیرد. همچنین برای نشان دادن عدم تاثیر تأخیر زمانی در پاسخ سیستم، شبیه سازی ها در هر مرحله برای دو مقدار $h = 0.5$ و $h = 3.5$ انجام می شود. ورودی اغتشاش را یک سیگنال سینوسی به صورت

$$w(t) = 0.1 \sin t$$

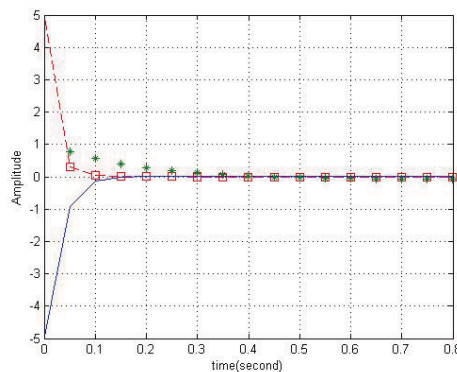
در نظر می گیریم. نتایج شبیه سازی برای $\gamma = 0.1$ در شکل های (۱) و (۲) نشان داده شده است. شکل (۱) مربوط به حالت بدون خطا و شکل (۲) مربوط به حالتی است که در سومین محرک خطا رخ داده است. همچنین به منظور مشاهده تضعیف صورت گرفته در اغتشاش وارد شده به سیستم، پاسخ حالت در مقیاسی بزرگتر در شکل (۳) رسم شده است. عدم تاثیر مقدار تأخیر در پایداری سیستم از شکل (۴) مشهود است. در این شکل پاسخ سیستم برای $h = 3.5$ رسم شده است. شکل های (۵) تا (۸) همین نتایج برای $\gamma = 3.1$ نشان می دهند. چنانکه از نتایج شبیه سازی ها مشاهده می شود، پایداری و رفتار مقاوم در مقابل اغتشاش، نامعینی و خرابی محرک تامین شده است.



شکل ۷ تضعیف اغتشاش برای حالتی که محرک سوم دچار خطا شود و

$$\gamma = 3.1, h = 0.5$$

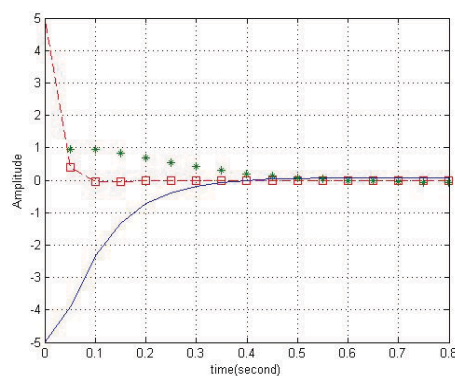
x_1 (نقطه)، x_2 (خط) و x_3 (ستاره)



شکل ۴ پاسخ حالت برای حالتی که محرک سوم دچار خطا شود، و

$$\gamma = 0.1, h = 3.5$$

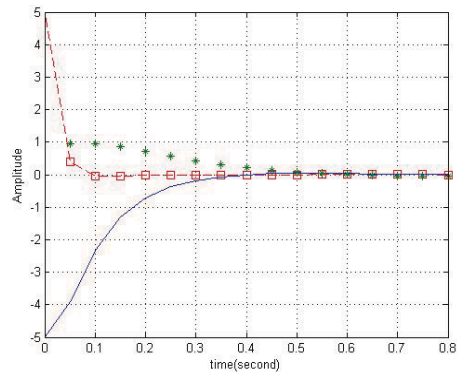
x_1 (مربع)، x_2 (خط) و x_3 (ستاره)



شکل ۸ تضعیف اغتشاش برای حالتی که محرک سوم دچار خطا شود و

$$\gamma = 3.1, h = 3.5$$

x_1 (مربع)، x_2 (خط) و x_3 (ستاره)



شکل ۵ پاسخ حالت برای حالتی که خطا رخ ندهد و با فرض $\gamma = 3.1, h = 0.5$

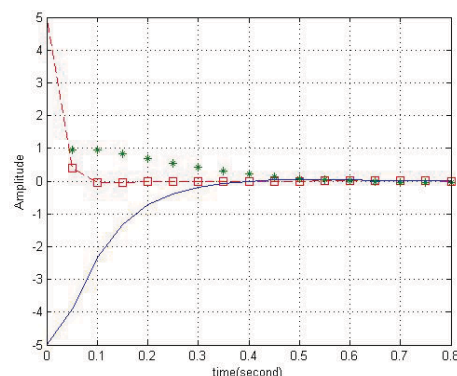
x_1 (مربع)، x_2 (خط) و x_3 (ستاره)

۵- نتیجه گیری

در این مقاله روشی برای طراحی مطمئن کنترل کننده H_∞ ارائه شده که پایداری و رفتار مناسب سیستم را هم در حالت عادی و هم در مواقع وقوع خطا در محرک های کنترل تضمین می کند. سیستم در نظر گرفته شده یک سیستم خطی با تأخیر زمانی و دارای نامعینی در تمام ماتریس های معادله حالت می باشد. بعلاوه روش ارائه شده مبتنی بر حل یک معادله ریکاتی توسعه یافته می باشد. در پایان کاربرد این روش با یک مثال عددی نشان داده شده است.

مراجع

[1] J.C., Doyle, K., Glover, P.P., Khargonekar and B.A., Francis, "State-space solutions to the standard



شکل ۶ پاسخ حالت برای حالتی که محرک سوم دچار خطا شود و

$$\gamma = 3.1, h = 0.5$$

x_1 (مربع)، x_2 (خط) و x_3 (ستاره)

- state delay and parameter uncertainty". *Automatica*, 32, 1183-1185, 1996
- [14] Vidyasagar, M. and Viswanadham, N., "Reliable stabilization using a multi controller configuration", *Automatica*, Vol. 21, No. 5, pp. 599-602, 1985.
- [15] Cho, Y.J., Bien, Z. and Kim, B.K., "Reliable control via additive redundant adaptive control" *In Proc. Amer. Contr. Conf.*, Pittsburgh, PA, 1989, pp. 1899-1904, 1989.
- [16] R. A. Date and J. H. Chow, "A reliable coordinated decentralized control system design," *in Proc. 28th Conf. Decision Contr.*, Tampa, FL, Dec, pp. 1295- 1300, 1989
- [17] M., Mariton and P., "Bertrand, Improved multiplex control systems: Dynamic reliability and stochastic optimality", *Int. J. Control*, 44,219-234, 1986
- [18] D.D., Siljak, "Reliable control using multiple control systems". *Int. J. Control*, 31, 303-329, 1980
- [19] Joshi, S.M., "Failure-accommodating control of large flexible space- craft" *In Proc. 986 Amer. Contr. Conf.*, Seattle, WA, pp. 156-161, 1986
- [20] R.J., Veillette, J.V., "Medanic and W.R., Perkins, "Design of reliable control systems". *IEEE Trans. Automat. Control*, 37,290-304, 1992
- [21] Z.D., Wang, X., Sun, J.S., Sun and Z.Q., "Wang Robust fault-tolerant control and design for uncertain linear systems", *Acta Aeronautica et Astronautica*, 17,112-115, 1996
- [22] Abootalebi, A., Hosseinnia, S. and Shekholeslam, F., "Reliable H_∞ control systems with uncertainties in all system matrices": An LMI approach", *IEEE Proc. Conf. on Control Application*, Canada, Aug 28-31, 2005.
- [23] J., Ge, P.M., Frank and C.-P., "Lin. Design of reliable controllers for state delayed systems", *European Journal of Control*, 2, 239-248, 1996
- [24] Z.D., Wang and H.Qiao, " H_∞ Reliable control of uncertain state delayed systems", *Journal of dynamic and control systems*, 55-76, 2004
- [25] R.J., Veillette. "Reliable state feedback and reliable observers". *Proc. 31st IEEE Conf. on Decision and Control*. Tucson, AZ, 2898-2903, 1992
- H_2 and H_∞ control problems", *IEEE Trans. Automat. Control*, 34,831-847, 1989.
- [2] Gu, K., " H_∞ control of systems under norm bounded uncertainties in all system matrices", *IEEE Trans. Automat. Control*, pp. 1320-1322, 1994.
- [3] Khargonekar, P.P., Petersen, I.R. and Zhou, K. "Robust stabilization of uncertain linear systems: quadratic stabilization and H_∞ control theory", *IEEE Trans. Automat. Control*, pp. 356-361, 1990.
- [4] De Souza, C.E., Fu, M. and Xie, L., " H_∞ analysis and synthesis of discrete-time systems with time-varying uncertainty", *IEEE Trans. Automat. Control*, pp. 459-462, 1993.
- [5] Xie, L., "Output feedback H_∞ control of systems with parameter uncertainty", *Int. J. Cont.*, 63, pp. 741-750, 1996.
- [6] Xie, L. and Soh, Y.C., "Robust control of linear systems with generalized positive real uncertainty." *Automatica*, 33, pp. 963-967, 1997.
- [7] Xie, L. and de Souza, C.E., "Robust H_∞ control for a class of uncertain linear time-invariant systems", *IEEE Proc. Control Theory, Appl.*, 138, pp. 479-483, 1991.
- [8] Xie, L. and de Souza, C.E., "Robust H_∞ control for linear systems with norm-bounded time varying uncertainty", *IEEE Trans. Automat. Control*, 37, pp. 1188-1191, 1992.
- [9] H.H., Choi and M.J., Chung, "Observer-based H_∞ controller design for state delayed linear systems", *Automatica*, 32, 1073-1075, 1996
- [10] H.H., Choi and M.J., Chung, "An LMI approach to H_∞ controller design for linear time delay systems", *Automatica*, 33, 737-739, 1997
- [11] C., Foias, A., Tannenbaum and G., Zames. "On the H_∞ -optimal sensitivity problem for systems with delays". *SIAM J. Control Optim.*, 25, 686-705, 1987
- [12] J.H., Lee, S.W., Kim and W.H., Kwon, "Memoryless H_∞ controllers for state delayed systems", *IEEE Trans. Automat. Control*. 39,159-162, 1994
- [13] J., Ge, P.M., Frank and C.-F., Lin, "Robust H_∞ state feedback control for linear systems with

ارائه روشی جدید در طراحی کنترلگر تطبیقی مقاوم برای سیستم های تاخیردار

سید کمال حسینی ثانی^۱، حمید رضا مومنی^۲

^۱ دکتری کنترل دانشگاه تربیت مدرس، Kamal_hosseini@yahoo.com

^۲ دانشیار گروه کنترل دانشگاه تربیت مدرس، Momeni_h@modares.ac.ir

چکیده: در این مقاله روش جدیدی برای کنترل سیستم های خطی تاخیردار ارائه شده است. این روش کنترلی، امکان طراحی کنترلگر تطبیقی پایدار با عملکرد مطلوب برای رنج مورد نظر از تاخیر زمانی سیستم را ممکن می سازد. به همین منظور ابتدا توسط جبران ساز فیدفوروارد، سیستم تاخیردار تبدیل به یک سیستم غیر فعال می شود و سپس با طراحی کنترلگر تطبیقی مدل مرجع جدید، کارایی مطلوب برای سیستم حاصل می گردد. به منظور مقایسه روش ارائه شده با روش کنترل ردیاب مولد فرمان، هر دو کنترلگر بر روی یک سیستم حرکت از راه دور نمونه، شبیه سازی شده اند.

واژه های کلیدی: تاخیر زمانی، کنترلگر تطبیقی مدل مرجع، جبران ساز فیدفوروارد، کنترلگر ردیاب مولد فرمان، سیستم غیر فعال، تابع تبدیل همیشه اکیدا حقیقی مثبت، سیستم حرکت از راه دور.

Abstract: In this paper, a new adaptive controller is proposed to ensure suitable stability and performance of the system with time delay while a wide range of time delay is considered. For this means, a feedforward compensator is designed to make the system passive first and then a new model reference adaptive controller (MRAC) is designed to have a good performance. Good stability and force tracking are the achievement of this controller. A command generator tracker (CGT) is designed for a sample teleoperation system and the results are compared with the proposed system.

Keywords: Time Delay, Model Reference Adaptive Controller, Feedforward Compensator, Command Generator Tracker, Passive System, Teleoperation System.

اندرسون و اسپانگ [۲] به عنوان پیشگامان این علم با ارائه مدل غیر فعال^۳ فعال^۳ برای کانال انتقال، پایداری سیستم را برای تاخیرهای بزرگتر کانال انتقال بهبود دادند ولی عملکرد سیستم چندان مناسب نبود. ماریوارن و لیرافان [۳] از تکنیک بهره متغیر، جهت غیر فعال کردن کانال انتقال استفاده کردند و پایداری سیستم را بهبود دادند. ریو و هانافورد [۴] استفاده از کنترلگر و رویتر جهت غیر فعال کردن سیستم را مطرح کردند. وظیفه رویتر، گزارش سیگنال های مورد نیاز برای بررسی غیر فعال بودن است در حالیکه کنترلگر، باعث غیر فعال کردن کل سیستم می شود. پاسخ های شبیه سازی نشان می دهند که رویت کردن سیگنال ها به صورت زمان واقعی برای نرم افزار مشکل است و نیاز به تحقیق بیشتر دارد. اسلوتین و نایمیر [۵]، از متغیرهای موج برای غیر فعال کردن کانال انتقال استفاده کردند و در سال ۱۹۹۷ [۶]، با بکار گیری فیلتر در مسیر متغیرهای موج و ارسال انتگرال متغیرهای موج به همراه متغیرهای موج، خطای ردیابی سیگنالهای موقعیت و سرعت کاهش داده شد. یک سال بعد توجه به این

۱- مقدمه

وجود تاخیر زمانی در اکثر سیستمها، باعث ناپایداری و عملکرد نامطلوب آنها می شود و به همین منظور محققین بسیاری روی آن کار کرده اند. سیستم حرکت از راه دور که در دهه اخیر کاربرد فراوانی پیدا کرده است نیز به دلیل تاخیر کانال انتقال از جمله سیستمهای تاخیردار می باشد. با بررسی آزمایشات انجام شده [۱] مشاهده می شود که زمان سیکل^۱ (زمان بین ارسال و دریافت سیگنال) برای سیستمهای روی مدار کوتاه زمین^۲ حداقل ۰/۴ ثانیه است، در حالیکه برای سیستمهای در سطح کره ماه حدود ۳ ثانیه می باشد. این مقادیر با لحاظ کردن زمان پردازش کامپیوترها روی ماهواره و در سطح زمین به حدود ۶ ثانیه می رسد. به همین منظور در این مقاله با در نظر گرفتن بیشترین زمان تاخیر ۷ ثانیه برای کانال انتقال، روشی جهت کنترل اینگونه سیستمها ارائه شده است.

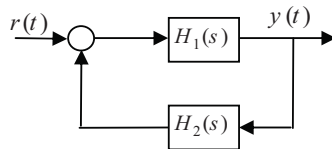
^۳ Passive Model

^۱ Cycle Time

^۲ Low earth Orbit

با مراجعه به قضیه ۱ مشاهده می شود اگر سیستم تحت کنترل، اکیدا غیر فعال خروجی^۲ باشد، کل سیستم حلقه بسته پایدار ورودی محدود-خروجی محدود^۳ خواهد بود.

قضیه ۱: اگر در سیستم حلقه بسته در شکل (۱)، H_1 اکیدا خروجی غیر فعال و H_2 غیر فعال باشند، در اینصورت سیستم حلقه بسته، پایدار ورودی محدود-خروجی محدود خواهد بود. [۱۲]



شکل (۱) یک سیستم حلقه بسته

با توجه به این قضیه و به دلیل اینکه کنترلگر H_2 غیر فعال است، به منظور پایدار سازی سیستم، لازم است که فرآیند را اکیدا غیر فعال خروجی نامییم. به همین منظور از جبران ساز فیدفوروارد بر اساس قضایای زیر استفاده می کنیم.

قضیه ۲: اگر $G(s) = G_1(s) + G_2(s)e^{-\delta s}$ اکیدا حقیقی مثبت باشد، آنگاه $G(s)$ اکیدا غیر فعال خروجی خواهد بود.

اثبات: ضمیمه ۱.

حال براساس قضیه فوق می توان به منظور اکیدا غیر فعال سازی فرآیند، شرط اکیدا حقیقی مثبت را مطابق با قضیه زیر بکار برد.

قضیه ۳: اگر جبران ساز فیدفوروارد $H(s)$ بر طبق شرایط زیر طراحی شود، آنگاه فرآیند اضافه شده^۴ $G_a(s) = G_p(s) + H(s)$ با انحرافات^۵ فرآیند، همیشه اکیدا حقیقی مثبت^۶ (ASPR) خواهد بود.

$$(G_p(s) = G(s)e^{-\delta s})$$

(۱) $H(s)$ پایدار باشد و درجه نسبی آن یک باشد.

(۲) فرآیند اضافه شده نامی $G_0(s) + H(s)$ همیشه اکیدا حقیقی مثبت باشد.

که در آن $\|\tilde{\Delta}(s)\|_{\infty} < 1$ و $\tilde{\Delta}(s) \in RH_{\infty}$ (۳)

$$\tilde{\Delta}(s) = \frac{G_0(s)W(s)}{G_0(s) + H(s)}$$

اثبات: مرجع [۱۱].

نکته که خراب شدن سیگنال موج به دلیل تغییرات زمان تاخیر کانال باعث خطای ردیابی و عدم برقراری غیر فعال بودن سیستم می شود آنها [۷] به جای انتقال سیگنال و سپس انتگرالگیری در سمت گیرنده، ابتدا انتگرال موج و انرژی را گرفته و سپس آن را همراه متغیرهای موج ارسال کردند و در گیرنده از فیلتر باز ساز^۱ استفاده نمودند و در نتیجه غیر فعال بودن سیستم مستقل از تاخیر کانال خواهد شد.

یدا و یاشیکوا [۸] در سال ۲۰۰۴ بحث بکارگیری فیلتر به منظور بهبود پاسخ سیستم در مقابل تاخیر زمانی کانال انتقال را مطرح کردند. لا و همکارانش [۹] در سال ۲۰۰۴ به صورت زمان واقعی توسط روش حداقل مربعات بازگشتی امیدانسان سایت دور را اندازه گرفتند و با بکارگیری روش جدیدی در انعکاس نیرو، انرژی مورد نیاز جهت انجام کار را برای اپراتور کاهش دادند. حسینی و همکاران [۱۰] بحث کنترلگر تطبیقی مدل مرجع را مطرح کردند که باعث بهبود پاسخ سیستم در مقابل تاخیرهای بزرگ کانال انتقال و بهبود عملکرد سیستم می شد. طراحی کنترلگر مدل مرجع با تخمین تاخیر زمانی و پیش بینی خروجی از جمله کارهای دیگری است که توسط این گروه انجام شده است.

در این مقاله سعی شده کنترلگری طراحی شود که با تضمین پایداری برای رنج مورد نظر از تاخیر زمانی سیستم، عملکرد مطلوبی برای سیستم بدست آید. در بخش ۲ ابتدا طراحی جبران ساز فیدفوروارد به منظور غیر فعال سازی فرآیند تحت کنترل مطرح شده است. در بخش ۳ روش جدیدی در طراحی کنترلگر تطبیقی مدل مرجع به منظور بهبود ردیابی سیستم ارائه شده است. در بخش ۴ نتایج شبیه سازی بر روی سیستم حرکت از راه دور مدل شده در [۱۰] آورده شده و دو روش با هم مقایسه شده اند. مقایسه این دو روش نشان می دهد که روش ارائه شده دارای پاسخ بهتری نسبت به روش CGT است و این محدودیت که روش CGT باید دارای جواب باشد را ندارد ولی اطمینان از مینیمم فاز بودن فرآیند خود یکی از محدودیتهای این روش است. بخش ۵ به نتیجه گیری از روش ارائه شده می پردازد.

۲- طراحی جبران ساز فیدفوروارد به منظور غیر فعال

سازی فرآیند تحت کنترل [۱۵]

^۱ Reconstruction Filter

^۲ Strictly Output Passive

^۳ Bounded Input Bounded output

^۴ Augmented System

^۵ Perturbation

^۶ Almost Strictly Positive Real

$$R u(t) = -S y(t) + T u_c(t)$$

که در آن R, S, T چند جمله ایهایی از درجه k, l و m هستند. از روابط (۷) و (۸) داریم:

$$R \frac{AA_1}{b_0(BA_1 + b_1B_1A)} y(t) = -S y(t) + T u_c(t)$$

برای سیستم می نیمم فاز می توان فرض کرد که همه صفرهای فرآیند با کنترلگر حذف شوند:

$$R = R_1(BA_1 + b_1B_1A)$$

بنابراین از روابط (۹) و (۱۰) داریم:

$$(A A_1 R_1 + b_0 S) y(t) = b_0 T u_c(t)$$

با توجه به رابطه فوق، مکان مطلوب جایابی قطب به صورت زیر قابل تعیین است.

$$AA_1 R_1 + b_0 S = A_0 A_m$$

که در آن A_0 قطبهای رویتگر و A_m قطبهای مطلوب مورد نظر هستند. با توجه با اینکه جبران ساز $H(s)$ به منظور غیرفعال سازی فرآیند طراحی شده، هدف از ردیابی کامل فرآیند آن است که خروجی فرآیند $(y_1(t))$ خروجی مرجع مطلوب را دنبال کند به عبارت دیگر داریم:

$$y_{m_1}(t) = y_1(t)$$

که در آن $y_{m_1}(t)$ خروجی مدل مرجع مطلوب است. از رابطه (۱۳) داریم:

$$A_0 A_m y_{m_1}(t) = A_0 A_m (y(t) - y_2(t)) =$$

$$A_0 A_m y(t) - A_0 A_m y_2(t)$$

با توجه به روابط (۱۲) و (۱۳):

$$y_{m_1}(t) = \frac{b_0 T}{A_0 A_m} u_c(t) - y_2(t) \Rightarrow$$

$$y_{m_1}(t) + y_2(t) = \frac{b_0 T}{A_0 A_m} u_c(t)$$

با معرفی

$$y_m(t) = y_{m_1}(t) + y_2(t) = \frac{b_0 T}{A_0 A_m} u_c(t)$$

مشاهده می شود که شرط ردیابی کامل آن است که خروجی $y(t)$ خروجی $y_m(t)$ را دنبال کند. با تعریف خطا به فرم زیر داریم:

$$e(t) = y(t) - y_m(t) \Rightarrow A_0 A_m e = A_0 A_m y - A_0 A_m y_m$$

با استفاده از روابط (۱۲) و (۱۶):

$$A_0 A_m e(t) = (AA_1 R_1 + b_0 S) y(t) - b_0 T u_c$$

با بکارگیری رابطه (۷) در رابطه (۱۸):

در روابط فوق $G_0(s)$ مقدار نامی تابع تبدیل $G(s)$ و $W(s)$ حد بالای $\Delta(s)$ است. $\Delta(s)$ نامعینی ضربی فرآیند است و براساس روابط زیر بدست می آید.

$$\Delta_d(s) = G(s) - G_0(s)$$

$$\Delta_m(s) = e^{-Ts} - 1$$

$$G_p(s) = (G_0(s) + \Delta_d(s))(I + \Delta_m(s))$$

$$\Delta(s) = \Delta_m(s) + G_0^{-1}(s) \Delta_d(s) (I + \Delta_m(s))$$

$$G_p(s) = G_0(s) (I + \Delta(s))$$

برای چک کردن شرط ۲ از قضیه ۳ می توان از قضیه زیر استفاده کرد.

قضیه ۴: تابع تبدیل $G(s)$ با ضرایب حقیقی، حقیقی مثبت است اگر و تنها اگر شرایط زیر برقرار باشند [۱۲].

(۱) $G(s)$ قطبی در سمت راست محور $j\omega$ نداشته باشد.

(۲) اگر تابع قطبهایی روی محور $j\omega$ یا در بینهایت داشته باشد، آنها ساده با مانده مثبت باشند.

$$Re(G(j\omega)) \geq 0 \quad (۳)$$

شرطهای ۱ و ۲ براحتی قابل تست هستند. برای تست شرط ۳ کافی است

که شرط زیر برای تابع $G(j\omega) = \frac{B(j\omega)}{A(j\omega)}$ چک شود.

$$f(j\omega) = Re(B(j\omega)A(-j\omega)) \geq 0 \quad \text{for } \omega \geq 0$$

بنابراین با توجه به قضیه های فوق، جبران ساز فیدفوروارد را به نحوی طراحی می کنیم که فرآیند اکیدا غیر فعال خروجی شود و با توجه به قضیه ۱، کنترلگر تطبیقی مدل مرجع طراحی شده یک سیستم حلقه بسته پایدار بدست خواهد داد. نکته مهمی که باید بدان توجه شود آن است که با موازی سازی جبران ساز فیدفوروارد با فرآیند، باید تمهیداتی در بلوک کنترل تطبیقی مدل مرجع داده شود که مساله ردیابی سیگنال خروجی حفظ گردد. به همین منظور در ادامه بحث، روش جدیدی در طراحی این کنترلگر ارائه شده است.

۳- روشی جدید در طراحی کنترلگر تطبیقی مدل مرجع با فیدبک از خروجی

فرض کنید فرآیند مورد نظر به فرم زیر باشد. (شکل ۲)

$$y(t) = y_1(t) + y_2(t) = \frac{b_0 B}{A} u(t) +$$

$$b_0 b_1 \frac{B_1}{A_1} u(t) \Rightarrow Ay(t) = b_0 B e^{-ds} u(t)$$

$$+ b_0 b_1 B_1 A / A_1 * u(t)$$

کنترلگر خطی را به صورت کلی زیر انتخاب می کنیم:

$$u = -\theta^T (P_1 \varphi)$$

در این صورت با استفاده از رابطه (۲۵) داریم:

$$e = b_0 (\varphi^T \dot{\theta} - \frac{1}{P_1} \theta^T (P_1 \varphi)) = b_0 *$$

$$(\varphi^T \dot{\theta} - \varphi^T \theta - \frac{1}{P_1} \theta^T (P_1 \varphi) + \varphi^T \theta)$$

ε و η را به صورت زیر تعریف می کنیم:

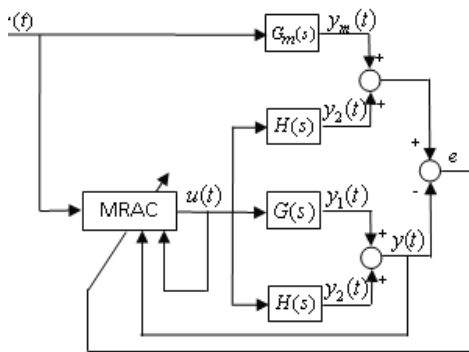
$$\eta = \frac{1}{P_1} \theta^T (P_1 \varphi) - \varphi^T \theta = -(\frac{1}{P_1} u + \varphi^T \theta)$$

$$\varepsilon = e_f + b_0 \eta = b_0 \varphi^T (\dot{\theta} - \theta)$$

سیگنال ε خطای افزوده نامیده می شود که به صورت زیر محاسبه می گردد.

$$\varepsilon = (y - y_m) + b_0 \eta$$

با توجه به طراحی فوق، بلوک دیاگرام کلی سیستم کنترل به فرم شکل (۲) قابل نمایش است.



شکل (۲) بلوک دیاگرام روش کنترلی ارائه شده

۴- نتایج شبیه سازی [۱۵]

بر اساس محاسبات انجام شده در [۱۰] (ضمیمه ۲)، تابع تبدیل سیستم مدل شده به صورت زیر است.

$$H_1(s) = \frac{\omega_s^2 (J_s s + b_s)}{(J_m s + b_m)(s^2 + 2\xi_s \omega_s s + \omega_s^2)} e^{-\delta s}$$

با توجه به روابط (۱) تا (۵) جبران ساز فیدفوروارد بفرم زیر قابل محاسبه است:

$$\Delta_m(s) = e^{-\delta s} - 1$$

$$\Delta_d(s) = 0$$

$$\Delta(s) = \Delta_m(s) = e^{-\delta s} - 1$$

$$G_0(s) = G(s)$$

$$A_0 A_m e(t) = A_1 R_1 b_0 \frac{B A_1 + b_1 B_1 A}{A_1} u(t)$$

$$+ b_0 S y(t) - b_0 T u_c(t)$$

$$\Rightarrow e(t) = \frac{b_0}{A_0 A_m} (R u(t) + S y(t) - T u_c(t))$$

ادامه مراحل طراحی مشابه فصل (۵) از مرجع [۱۲] مطابق فرمولهای زیر قابل انجام است.

چون ممکن است $\frac{b_0}{A_0 A_m}$ اکیدا حقیقی مثبت نباشد و بنابراین نتوان از

قانون لیاپونف استفاده کرد، پس از مدل خطای زیر استفاده می کنیم:

$$e_f = b_0 (\frac{R}{P} u + \frac{S}{P} y - \frac{T}{P} u_c)$$

که در آن $P = P_1 P_2 = A_0 A_m$ چند جمله ای یکتا، پایدار و با درجه مشابه درجه R است در این صورت

$$\frac{R}{P} = \frac{1}{P_1} + \frac{R - P_2}{P}$$

بنابراین مدل خطا به فرم زیر قابل نمایش است.

$$e = b_0 (\frac{1}{P_1} u + \frac{R - P_2}{P} u + \frac{S}{P} y - \frac{T}{P} u_c)$$

بردار پارامترهای صحیح کنترلگر به فرم زیر قابل تعریف است.

$$\theta^0 = (r'_1 \dots r'_k s_0 \dots s_1 t_0 \dots t_m)^T$$

که در آن r'_i ها ضرایب چند جمله ای $R - P_2$ هستند. بردار φ شامل ورودی، خروجی و سیگنال فرمان به صورت زیر قابل تعریف است.

$$\varphi^T = (\frac{\rho^{k-1}}{P(\rho)} u \dots \frac{1}{P(\rho)} u \quad \frac{\rho^l}{P(\rho)} y \dots \frac{1}{P(\rho)} y \\ - \frac{\rho^m}{P(\rho)} u_c \dots - \frac{1}{P(\rho)} u_c)$$

بنابراین سیگنال خطا به فرم زیر قابل نمایش است.

$$e = b_0 (\frac{1}{P_1} u + \varphi^T \dot{\theta})$$

در حالت نامی که پارامترهای کنترلگر معلوم هستند داریم:

$$u = -p_1 (\varphi^T \dot{\theta}) = -p_1 ((\dot{\theta}^T) \varphi) = -(\dot{\theta}^T) (p_1 \varphi)$$

با فرض اینکه θ بردار پارامترهای قابل تنظیم کنترلگر باشد قانون فیدبک زیر مدل خطای مطلوب را بدست می دهد.

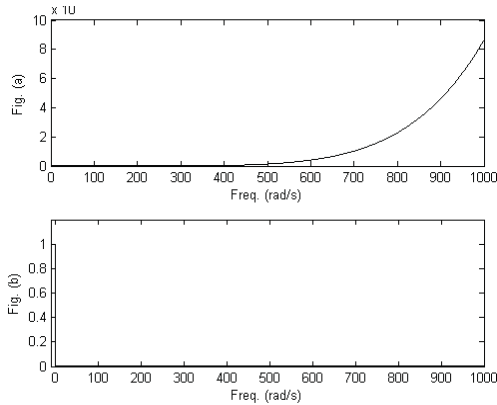
$$u = -p_1 (\varphi^T \theta)$$

با توجه به اینکه این قانون کنترلی درحالی که درجه p_1 بزرگتر از یک

باشد قابل تحقق نیست (زیرا جمله $p_1 (\varphi^T \theta)$ شامل مشتقات پارامترها

خواهد بود)، از قانون کنترلی قابل تحقق زیر استفاده می شود.

این قضیه را برقرار می سازد. همچنین شرط (۳) این قضیه هم با رسم $\|\tilde{\Delta}(S)\|_{\infty}$ مطابق شکل (۴b) برقرار است. برای بررسی شرط (۲) قضیه (۳) از قضیه (۴) کمک گرفته شده و رابطه (۲-۶)، وقتی که تابع تبدیل $G_0(S)+H(S)$ است، در شکل (۴a) رسم شده است. مشاهده می شود که این مقدار برای تمامی فرکانسها مقدار مثبتی است و بنابراین شرط (۲) قضیه (۳) هم برقرار است.



شکل (۴)

a: برقراری شرط (۲-۶) برای بررسی ASPR بودن $G_0(S)+H(S)$
 $\|\tilde{\Delta}(S)\|_{\infty}$

بنابراین جبران ساز $H(S)$ انتخاب شده، سیستم حرکت از راه دور را برای تاخیر زمانی رفت و برگشت کانال تا ۷ ثانیه، ASPR و بر اساس قضیه (۲) غیر فعال می سازد.

با انتخاب $H(S)$ ، تابع تبدیل کل سیستم از درجه ۴ است و بر اساس شرایط مناسب بدون^۲ کنترلگر (رابطه ۴۱ و ۴۲) و شرط سازگاری^۳ (رابطه ۴۳ تا ۴۶)، درجه چند جمله ایهای T, S, R به صورت زیر بدست خواهد آمد.

$$\deg R = 3, \quad \deg S = 3, \quad \deg T = 0$$

شرط مناسب بدون:

$$\deg S \leq \deg R$$

$$\deg T \leq \deg R$$

شرط سازگاری:

$$\deg A_{mod} = \deg A,$$

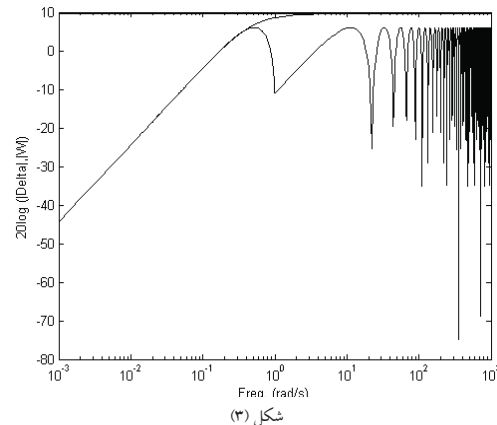
$$\deg B_{mod} = \deg B,$$

$$\deg A_0 = \deg A - \deg B - 1,$$

با رسم اندازه $\Delta(S)$ می توان تابع تبدیل $W(S)$ را به فرم زیر، به عنوان حد بالای اندازه $\Delta(S)$ بدست آورد.

$$W(s) = \frac{3s}{s + 0.4}$$

در شکل (۳) اندازه $\Delta(S)$ و $W(S)$ رسم شده است.



شکل (۳)

a: $|\Delta(S)|$, b: $|W(s)| = \left| \frac{3s}{s+0.4} \right|$ c:

$$|W(s)| = 20 \log(3)$$

در محاسبه $W(S)$ مقدار $d = T_1 + T_2$ را برابر ۷ ثانیه انتخاب کرده ایم که مقدار بزرگی برای تاخیر یک سیستم حرکت از راه دور است. برای $W(S)$ مقدار دیگری از جمله $W(S) = 3$ نیز قابل انتخاب است که کمی نسبت به مقدار قبلی محافظه کارانه^۱ می باشد.

اکنون برای محاسبه $H(S)$ از حداقل کردن تابع $\|\tilde{\Delta}(S)\|_{\infty} < 1$ استفاده می کنیم.

$$\tilde{\Delta}(S) = \frac{\frac{s+3}{(170s+2)(s^2+1.4s+1)} * 3}{\frac{s+3}{(170s+2)(s^2+1.4s+1)} + H(s)}$$

تابع تبدیل $H(S)$ را سریعتر از فرآیند انتخاب می کنیم و به منظور ساده سازی محاسبات $H(S)$ از درجه یک به فرم زیر انتخاب شده است:

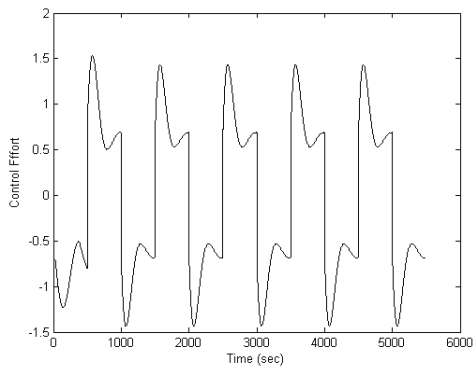
$$H(s) = \frac{b_0}{(s+10)}$$

با توجه به اینکه $\|\tilde{\Delta}(S)\|_{\infty}$ مقدار بهره ماکزیمم $\tilde{\Delta}(S)$ را نمایش می دهد، می توان محاسبات را به صورت تحلیلی نیز انجام داد. نتایج محاسبات مقدار $b_0 = 30$ را بدست خواهد داد.

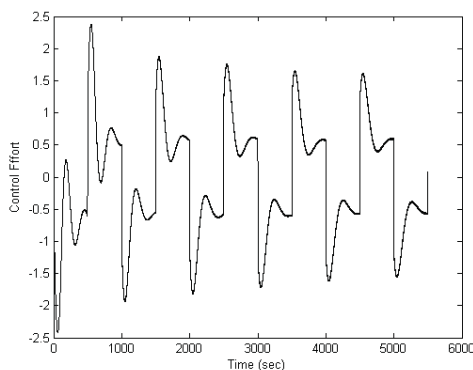
برای اینکه نشان دهیم $H(S)$ طراحی شده شرط ASPR را برقرار می سازد با رجوع به قضیه (۳) مشاهده می شود که $H(S)$ انتخابی شرط (۱)

² Proper
³ Compatibility

¹ Conservative



شکل (۷) سیگنال کنترلی در روش جدید



شکل (۸) سیگنال کنترلی در روش کنترل CGT

نتایج شبیه سازی نشان می دهد که روش کنترلی ارائه شده دارای ردیابی مناسبی برای سیگنالهای سرعت و نیرو می باشد و سیگنال کنترلی مورد نیاز نیز نسبت به روش CGT مقدار کمتری دارد.

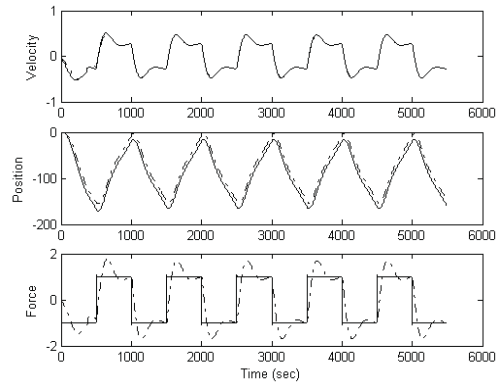
۵- نتیجه گیری

در این مقاله کنترلگر تطبیقی مدل مرجع مناسبی برای سیستم های خطی تاخیردار ارائه شده است. قابلیت پایدارسازی و عملکرد مطلوب این کنترلگر برای محدوده مورد نظر از تغییرات پارامترهای فرآیند و تغییرات زمان تاخیر سیستم از جمله مزایای آن است.

مقایسه این روش با روش CGT نشان نشان می دهد که روش کنترلی ارائه شده دارای پاسخ بهتری نسبت به روش CGT است. مزیت این روش نسبت به روش CGT آن است که اولاً در روش CGT باید برای مسئله مورد نظر جواب وجود داشته باشد (رابطه (۲۱) مقاله [۱۵]). ثانیاً در روش CGT بهره های تطبیقی در محاسبه سیگنال کنترلی به صورت PI انتخاب می شوند در حالی که در روش ارائه شده کنترلگر می تواند به صورت دلخواه و کلی تر $(Ru=Tu_c - Sy)$ انتخاب گردد که

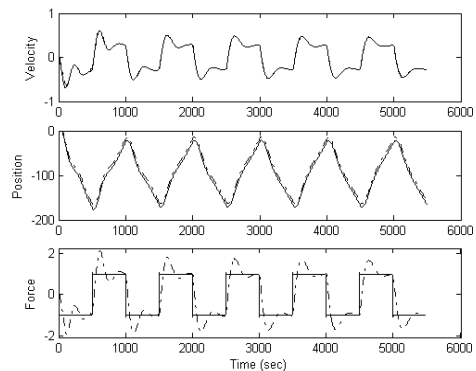
$$A R_1 + \beta S = A_0 A_{mod}$$

در ادامه این بخش، نتایج شبیه سازی برای ورودی پالس مربعی به ربات فرمانده و به ازای تاخیر بزرگ ۵ ثانیه برای دو حالت کنترلگر مدل مرجع ارائه شده و کنترلگر CGT [۱۵]، آورده شده است.



شکل (۵) مقایسه سیگنالهای سرعت، موقعیت و نیرو برای تاخیر زمانی ۵ ثانیه و با روش

کنترلی ارائه شده



شکل (۶) مقایسه سیگنالهای سرعت، موقعیت و نیرو برای تاخیر زمانی ۵ ثانیه و با روش

کنترل CGT

همانطور که از شکل (۵) و (۶) مشاهده می شود با وجود تاخیر زمانی بزرگ کانال انتقال، فرآیند تحت کنترل پایدار مانده و دارای ردیابی مناسبی است.

در شکل های (۷) و (۸) سیگنال کنترلی برای دو روش فوق رسم شده اند.

$$\tau_m = -b_m \dot{u}_m - F_{md}$$

$$\tau_s = -b_s \dot{u}_s + F_s$$

که در آن u_m و u_s به ترتیب سرعت رباتهای فرمانده و فرمانبر، J_m و J_s اینرسی، b_m و b_s دمپینگ، F_h نیروی اعمالی از دست اپراتور به ربات فرمانده، τ_m و τ_s گشتاور موتورها، F_{md} نیروی برگشتی از سایت فرمانبر، F_s گشتاور هماهنگ ساز^۱ و F_e گشتاور محیط است. گشتاور F_s معمولاً جهت تطبیق سیگنالهای سرعت فرمانده و فرمانبر و توسط کنترلگر اعمال می شود. با توجه به روابط فوق تابع تبدیل رباتهای فرمانده و فرمانبر قابل محاسبه هستند.

$$H_m(s) = \frac{u_m}{F_h - F_{md}} = \frac{1}{J_m s + b_m}$$

$$H_s(s) = \frac{u_s}{F_s - F_e} = \frac{1}{J_s s + b_s}$$

کانال انتقال نیز می تواند با یک تاخیر خالص مدل گردد. بنابراین داریم:

$$u_{sd}(s) = u_m(t - T_1)$$

$$F_{md}(t) = F_s(t - T_2)$$

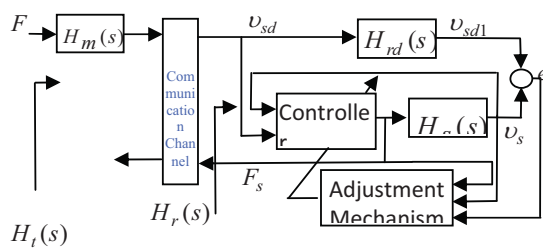
که T_1 و T_2 تاخیر کانال رفت و برگشت است و u_{sd} سرعت اعمال شده به سایت فرمانبر و F_{md} نیروی برگشتی از سایت فرمانبر است.

بنابراین داریم:

$$u_{sd}(s) = e^{-T_1 s} \cdot u_m(s)$$

$$F_{md}(s) = e^{-T_2 s} \cdot u_m(s)$$

با فرض $F_e = 0$ و با توجه به روابط فوق و بکارگیری مدل مرجع در سمت فرمانبر، دیاگرام کلی سیستم مطابق شکل (۹) ارائه می گردد.



شکل (۹) بلوک دیاگرام سایت فرمانده

نوع PI را هم می تواند شامل باشد. بنابراین انتخاب مناسب تر کنترلگر پاسخ بهتری را برای سیستم نسبت به روش CGT به همراه خواهد داشت. ولی باید توجه کرد که روش انتخابی برای سیستمهای مینیم فاز قابل اعمال است و برای سیستمهای غیر مینیم فاز باید از روشهای کنترلی مربوط به آنها استفاده نمود.

ضمیمه ۱ (اثبات قضیه ۲)

از قضیه بار سوال داریم:

$$\langle y/u \rangle = \int_0^{\infty} y(t)u(t)dt =$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} Y(i\omega)U(-i\omega)d\omega =$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} G(i\omega)U(i\omega)U(-i\omega)d\omega =$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} G_1(i\omega)U(i\omega)U(-i\omega)d\omega +$$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} G_2(i\omega)e^{-T_1\omega}U(i\omega)U(-i\omega)d\omega =$$

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \text{Re}(G_1(i\omega))U(i\omega)U(-i\omega)d\omega +$$

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \text{Re}(G_2(i\omega)e^{-T_1\omega})U(i\omega)U(-i\omega)d\omega =$$

$$\frac{1}{\pi} \int_{-\infty}^{\infty} \text{Re}(G_2(i\omega))U(i\omega)U^*(-i\omega)d\omega$$

حال اگر $G(s)$ اکیدا حقیقی مثبت باشد:

$$\text{Re}(G(i\omega)) > \varepsilon \Rightarrow \langle y/u \rangle^2 =$$

$$\frac{\varepsilon}{\pi} \int_0^{\infty} |G(i\omega)|^2 U U^* d\omega$$

$$\Rightarrow \langle y/u \rangle^2 \geq \varepsilon_1 \|G(i\omega)\|_2^2$$

یعنی $G(s)$ اکیدا غیر فعال خروجی است.

ضمیمه ۲ (مدل سازی سیستم حرکت از راه دور نمونه)

مدلسازی سیستم حرکت از راه دور نمونه به صورت کامل در [۱۰] توضیح داده شده است. در این بخش نیز به منظور دستیابی سریعتر خواننده، مراحل مدلسازی آورده شده است. سیستم مورد نظر، یک سیستم از نوع ردیابی نیرو/ سرعت است و با توجه به اینکه در اکثر مقالات از ربات یک درجه آزادی برای سایت فرمانده و فرمانبر استفاده شده ما نیز به منظور مقایسه نتایج از همان مدل استفاده می کنیم. معادلات (۴۷) تا (۵۰) دینامیک ربات فرمانده و فرمانبر را نشان می دهند.

$$J_m \dot{u}_m = F_h + \tau_m$$

$$J_s \dot{u}_s = -F_e + \tau_s$$

^۱ Coordinating Torque

- [4] J. Ryu, D. Kwon, B. Hannaford, "Stable Teleoperation with Time Domain Passivity Control", *proc. IEEE int conf. on Robotics & Automation*, May 2002
- [5] G. Niemeyer and J.-J. E. Slotine, "Stable adaptive teleoperation," *IEEE J. of Oceanic Engineering*, vol. 16, no.1, Jan. 1991.
- [6] G. Niemeyer, J.-J. E. Slotine, "Using Wave Variables For System Analysis And Robot Control," *Proc. IEEE Int. Conf. On Robotics & Automation*, pp. 1619-1625, 1997.
- [7] G. Niemeyer and J.-J. E. Slotine, "Towards Force-Reflecting Teleoperation over the Internet," *Proc. IEEE Int. Conf. on Robotics & Automation*, Leuven, Belgium, pp. 1909-1915, 1998.
- [8] J. Ueda and T. Yoshikawa "Force-Reflecting Bilateral Teleoperation with Time Delay by Signal Filtering", *IEEE Transactions on Robotics & Automation*, vol. 20, no. 3, June 2004.
- [9] L. J. Love and W. J. Book "Force Reflecting Teleoperation With Adaptive Impedance Control", *IEEE Transactions on Systems, Man, and Cybernetics—Part*, vol. 34, no. 1, February 2004.
- [10] K. H. Sunny, H. Momeni and F. Janabi-Sharifi, "Adaptive teleoperation systems design," *Proc. IEEE Int. Conf. Control Applications, CCA2005*, Toronto, Canada, August 2005.
- [11] Kaufman, H. and G. Neat, "Asymptotically Stable MIMO Direct MRAC for Processes not Necessarily Satisfying a Positive Real Constraint," *Int. J. Control*, 58, pp. 1011-1031, 1993.
- [12] K. J. Astrom, *Adaptive Control System*, 1995, Addison-Wesley, New York, USA.
- [13] Broussard, J. and O. O'Brien, "Feedforward control to track the output of a forced model," *IEEE 17th Conference on Decision and Control*, San Diego, CA, vol. 42, pp. 1149-1155, 1979.
- [14] Kaufman, H., I. Bar-Kana, and K. Sobel. *Direct Adaptive Control Algorithms*, Springer-Verlag, 1998.
- [15] سید کمال حسینی ثانی، حمید رضا مومنی، "طراحی کنترلگر تطبیقی مقاوم برای سیستم های حرکت از راه دور"، *مجله کنترل و ابزار دقیق*، جلد ۱، شماره ۱، ۱۳۸۵.

در شکل فوق $H_{rd}(s)$ مدل مرجع مورد نظر در طراحی کنترلگر تطبیقی سایت فرمانبر است که برای آن یک تابع تبدیل درجه ۲ به فرم زیر در نظر گرفته شده است.

$$H_{rd}(s) = \frac{\omega_s^2}{s^2 + 2\zeta_s \omega_s s + \omega_s^2}$$

در صورت حصول ردیابی مناسب در سمت سایت فرمانبر داریم: $U_{sd1} = U_s$ و بنابراین تابع تبدیل قابل مشاهده از سمت سایت فرمانده به صورت زیر قابل محاسبه است.

$$H_r(s) = \frac{F_s(s)}{U_{sd}(s)} = \frac{F_s}{U_s} \cdot \frac{U_{sd1}}{U_{sd}} = H_s^{-1}(s) \cdot H_{rd}(s)$$

$$H_t(s) = \frac{F_{md}(s)}{F_h - F_{md}(s)} = H_m(s) \cdot e^{-T_1 s}$$

$$H_r(s) \cdot e^{-T_2 s} = \frac{H_m(s) \cdot H_{rd}(s)}{H_s(s)} e^{-ds}$$

که $d = T_1 + T_2$ و $H_t(s)$ تابع تبدیل سایت فرمانبر و $H_r(s)$ تابع تبدیل از دیدگاه سایت فرمانده است. با جایگزینی روابط (۵۱)، (۵۲) و (۵۳) در رابطه (۵۹) داریم:

$$H_t(s) = \frac{\omega_s^2 (J_s s + b_s)}{(J_m s + b_m)(s^2 + 2\zeta_s \omega_s s + \omega_s^2)} \cdot e^{-ds}$$

رابطه (۶۰) تابع تبدیل سیستم حرکت از راه دور را نشان می دهد.

مراجع

- [1] L. F. Penin and K. Matsumoto, "Teleoperation With Time Delay. A Survey and Its Use in Space Robotics," *Technical Report of National Aerospace Laboratory TR-1438T*, 2000.
- [2] R. J. Anderson, M. W. Spong, "Bilateral control of teleoperators with time delay," *IEEE Trans. Automatic Control*, vol. 34, no.5, May 1989, pp. 494-501.
- [3] S. Leeraphan, T. Marieewarn, and D. Laowattana, "Stable adaptive bilateral control of transparent teleoperation through time-varying delay," *Proc. IEEE/RSJ, Int. Conf. on Robots and systems*. Lausanne, Switzerland, pp. 2979-2984, October 2002.

تأثیر بردار برتر در دینامیک فرآیندهای QRDS و کاربرد آن

در کنترل فرآیندهای دارای تاخیر زمانی

منصور شیروانی^۱، منصوره اسماعیلی^۲^۱ دانشیار، دانشکده مهندسی شیمی، دانشگاه علم و صنعت ایران. shirvani.m@iust.ac.ir^۲ دانشجوی دکترا، دانشکده مهندسی شیمی، دانشگاه علم و صنعت ایران. mansoreh@iust.ac.ir

چکیده: در این مقاله با استفاده از مفهوم بردار برتر و تأثیر آن در دینامیک توابع تبدیل با ساختار Irrational، روشی جدید برای کنترل فرآیندهای دارای تاخیر زمانی ارائه می شود. در این روش در معادله مشخصه سیستم کنترل، تابعی درجه یک با اندازه برتر نسبت به جملات دیگر تابع حلقه باز سیستم کنترل، اضافه شده و بدین ترتیب رفتار دینامیکی سیستم که به دلیل وجود پارامتر تاخیر زمانی در حوزه فرکانسی بصورت تاخیر زمانی دار می باشد با تابعیت از رفتار بردار برتر تبدیل به رفتار بدون تاخیر زمانی خواهد شد و صفرهای سمت راست موجود در تابع حلقه باز سیستم کنترل، که به دلیل وجود پارامتر تاخیر زمانی در آن به وجود آمده اند، از بین رفته و سیگنال ورودی به کنترلر فاقد این صفرها خواهد شد. از قابلیت‌های مهم این روش عدم حساسیت آن به خطاهای موجود در مدل فرآیند و همچنین قابلیت کاربرد مستقیم در کنترل فرآیندهای پارامتر گسترده با ساختار QRDS می باشد.

واژه های کلیدی: جبران کننده تاخیر زمانی، بردار برتر، رفتار تاخیر زمانی دار، محدود کننده فاز، مقاوم بودن

Abstract: In this paper, a new method is proposed for controlling the processes including Time-Delay (TD). The method is based on the concept of dominant gain and its effect on the dynamics of irrational structured transfer function models. In this method, a minimum phase transfer function, which is almost required to be a first order one, is used for establishing the dominant gain requirement in control loop. This function is used as a secondary inner loop feedback from the controller output signal to the input such that establishment of the dominant gain constraint in the characteristic equation becomes possible. In this way, the RHP zeros of the open loop transfer function will be removed perfectly and the input signal to the controller will become free of the effect of such zeros. Among the important capabilities of this method is its much minor sensitivity to the model error and also its straight applicability for controlling QRDS processes.

Keywords: Dead-Time Compensator, Dominant Vector, Time-delayed Behavior-gain, Robustness

انجام شده است که در مرجع [۱] در مورد این روشها توضیحات مختصری به همراه ذکر مراجع آن داده شده است. آنچه در کنترل این سیستمها باید در نظر داشت این است که وجود پارامتر تاخیر زمانی در تابع حلقه باز سیستم کنترل ممکن است باعث به وجود آمدن صفرهای سمت راست شود که می تواند موجب ناپایداری سیستم کنترل شود، لذا در روشهای کنترل این سیستمها، حذف صفرهای سمت راست از تابع حلقه باز سیستم کنترل همواره مد نظر بوده است. از رایجترین روشهای کنترل سیستمهای دارای تاخیر زمانی می توان به جبران کننده های تاخیر زمانی "Dead-Time Compensators (DTCs)" اشاره کرد [۲]. این روشها بر اساس روش جبران کننده اسمیت بنا نهاده شده اند. روش اسمیت بر اساس حذف پارامتر تاخیر زمانی از تابع حلقه باز سیستم کنترل استوار

۱- مقدمه

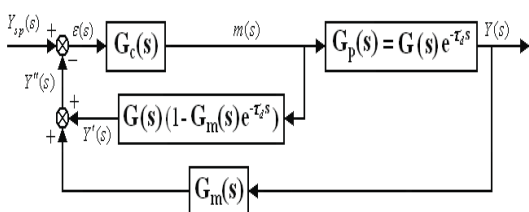
بسیاری از فرآیندهای موجود در صنایع شیمیایی با تاخیر زمانی همراهند. مدل این فرآیندها را پس از خطی سازی می توان به صورت یک چند جمله ای به همراه پارامتر تاخیر زمانی نمایش داد. کنترل این فرآیندها به دلیل وجود تاخیر زمانی در آنها با استفاده از کنترل کننده های مرسوم بسیار مشکل می باشد چرا که وجود تاخیر زمانی باعث کاهش کیفیت کنترل میشود و هر چه مقدار پارامتر آن بیشتر باشد تأثیر آن شدید تر است. به دلیل وجود این مشکلات از سالها قبل تلاشهای زیادی برای ابداع روشهای مختلف برای کنترل فرآیندهای دارای تاخیر زمانی

قسمت اول این مقاله نشان داده شد که پیش بینی رفتار دینامیکی مدل‌های QRDS در حوزه فرکانسی با استفاده از مفهوم بردار برتر در تطابق با تئوری مکان مجانبی صفرها می باشد.

در روش ابداع شده در این مقاله با اضافه کردن تابع انتقال درجه یک با اندازه (Gain) برتر نسبت به اندازه تابع حلقه باز (Open Loop Gain) سیستم کنترل دارای تأخیر زمانی، تأثیر جمله دارای تأخیر زمانی را در نمودار فاز تابع انتقال حلقه باز کاهش داده و بنابراین صفرهای واقع در سمت راست تابع حلقه باز سیستم کنترل را حذف می‌کنیم. بدین ترتیب، رفتار فرکانسی تابع تبدیل حلقه باز سیستم کنترل را از وضعیت رفتار تأخیر زمانی دار به وضعیت بدون تأخیر زمانی تغییر می‌دهیم. مزیت مهم این روش در سادگی استفاده و مقاومت بالای آن (Robustness) در مقابل خطای مدل می باشد. مزیت مهم دیگر این روش این است که میتوان آترو مستقیماً برای کنترل فرآیندهای با ساختار مدلی QRDS نیز بکار برد. در بخشهای دوم، سوم و چهارم مقاله در مورد ساختار و مشخصات سیستم کنترل مورد نظر بحث خواهد شد. در بخش پنجم پاسخهای شبیه سازی شده سیستم برای فرآیندهای مختلف ارائه می شود و در بخش ششم نیز نتیجه گیری کلی از مقاله ارائه خواهد شد.

۲- ساختار سیستم کنترل پیشنهادی

در کنترل فرآیندهای دارای تأخیر زمانی دسترسی به مدل فرآیند برای پیش بینی کردن سیگنالها از اهمیت ویژه ای برخوردار است. همانطور که قبلاً هم ذکر شد روش پیش بین کننده اسمیت یکی از روشهای رایج کنترل سیستمهای دارای تأخیر زمانی می باشد که بر اساس پیش بینی مدل فرآیند ابداع شده است. ساختار کنترل در این روش در شکل ۱ نشان داده شده است.



شکل ۱. حلقه کنترل پیش بین کننده اسمیت

در این روش، اسمیت از یک پیش بین کننده استفاده می کند. این پیش بین کننده سیگنال ورودی به کنترل کننده را طوری تغییر می دهد که پارامتر تأخیر زمانی در آن بطور کامل از بین برود که بدین ترتیب بر اساس سیگنال ورودی به کنترل می توان مقادیر بالاتری از پارامترها را برای کنترل انتخاب نمود و این باعث بهتر شدن پاسخ خروجی خواهد

است [۳]. هر چند بعدها مشخص شد که روش وی دچار مشکلاتی نظیر حساسیت زیاد نسبت به خطای مدل [۷-۴] و رفع نکردن افست در فرآیندهای انتگرالی به هنگام بروز آشفتگی می باشد [۱۸-۸]. برای رفع مشکلات و نارساییهای موجود در روش اسمیت تا سالهای اخیر محققان بسیاری بر روی آن کار کرده اند که نتیجه آن معرفی روشهای مختلف DTCS بوده است [۱۸-۵ و ۶]. همه آنها بر اساس حذف پارامتر تأخیر زمانی از تابع حلقه باز سیستم کنترل می باشند. اما بزرگترین عیب کلیه روشهای مبتنی بر DTCS کاربرد نداشتن این روشها برای کنترل فرآیندهایی است که پارامتر تأخیر زمانی از تابع فرآیند آنها قابل فاکتور گیری و تفکیک نمی باشد، یعنی فرآیند را نمی توان به دو عبارت که یکی دارای تأخیر زمانی و دیگری بدون تأخیر زمانی باشد تقسیم کرد. در اصطلاح به این توابع انتقال، Irrational و به این سیستمها "Quasi-Rational Distributed System (QRDS)" می گویند. رفتار دینامیکی این سیستمها در مراجع مختلفی مورد بررسی قرار گرفته است [۲۴-۲۰]. در [۲۱ و ۲۰] با بررسی رفتار دینامیکی آنها و محاسبه صفرهای سمت راست تابع حلقه باز سیستم کنترل، روشی برای کنترل آنها ارائه شده است. در این روش با الهام از روش اسمیت و توجه به اینکه در این روش با حذف پارامتر تأخیر زمانی از تابع حلقه باز سیستم کنترل و سیگنال ورودی به کنترل در اصل صفرهای سمت راست حذف شده اند پیش بین کننده ای طراحی شده است که صفرهای سمت راست تابع حلقه باز سیستم کنترل را حذف کرده و مانع ارسال سیگنال حاوی این صفرها به کنترلر می شود. بدین ترتیب بدون اینکه سیستم ناپایدار شود می توان بهره بالاتری برای کنترلر انتخاب کرد و پاسخ آن را نسبت به یک حلقه پس خور معمولی بهبود بخشید. در سالهای اخیر نیز روشهای مشابهی که بر پایه حذف صفرهای سمت راست از تابع حلقه باز سیستم کنترل استوارند، برای بهبود پاسخ این سیستمها ارائه شده است [۲۷-۲۵]. اما بزرگترین عیب این روشها این است که محاسبه دقیق و آسان صفرهای سمت راست امکانپذیر نمی باشد و این موضوع حتی تا سالهای اخیر موضوع تحقیق محققان بوده است [۲۹ و ۲۸]. در حقیقت باید گفت تمامی راه حل‌های ارائه شده عددی است و این صفرها را تنها در محدوده خاصی از فرکانسها ارائه میدهد. در این مقاله بر اساس شناخت رفتار دینامیکی سیستمهای QRDS که در مقاله اول بطور مفصل راجع به آن بحث شد، روشی نوین برای کنترل فرآیندهای دارای تأخیر زمانی ارائه شده است. همانطور که در بخش اول مقاله ذکر شد رفتار سیستمهای QRDS در حوزه پاسخ فرکانسی همواره و در هر بازه فرکانسی تحت تأثیر بردار با اندازه بزرگتر که اصطلاحاً در اینجا بردار برتر نامیده می شود قرار دارد. همچنین در

بهره المان اندازه گیر می باشد. از معادله ۴ نیز می توان مقدار افست به هنگام ایجاد آشفتگی را به دست آورد که این مقدار برابر صفر می باشد. از روابط فوق معادله مشخصه بصورت زیر می باشد:

$$1 + G_{mb}(s) + G_c(s)G_f(s)G_p(s)G_m(s) = 0 \quad (5)$$

تابع G_{mb} در رابطه (۵) تابع انتقال با اندازه برتر (Dominant Gain) نسبت به ترم سوم معادله که تابع انتقال حلقه باز سیستم کنترل ساده می باشد، در نظر گرفته می شود. در اینجا ما با افزودن آن به معادله مشخصه در شرایطی که دارای بهره برتری باشد، رفتار حلقه باز سیستم کنترل را تابع آن کرده و صفراهای سمت راست موجود در حلقه باز را بکمک آن مرتفع می نماییم. این تابع در اصل نقش پیش بین کننده اسمیت را بازی می کند و در اینجا به نام پیش بین کننده Model Bypass Phase Limiter (MBPL) نامگذاری شده است. این نامگذاری انعکاس دهنده این حقیقت است که پیش بین کننده فوق قادر به دور زدن اثرات تاخیر زمانی در تابع تبدیل حلقه باز سیستم بوده و بواسطه برتری اندازه قادر است که نمودار فاز حلقه باز سیستم کنترل را محدود نماید. همانطور که می دانیم یک حلقه کنترل ساده با یک تابع انتقال حلقه باز درجه یک همواره پایدار است. لذا بنظر می رسد انتخاب یک تابع درجه یک برای G_{mb} با اندازه برتر نسبت به ترم دیگر موجود در معادله مشخصه می تواند شرایط خوبی را برای پایداری به وجود آورد. بنابراین در مثالهای مطرح شده در این مقاله، یک G_{mb} یک تابع درجه یک انتخاب شده است.

۳- بررسی سیگنال ورودی به کنترل کننده در روشهای مختلف

در اینجا سیگنالهای ورودی به کنترل کننده در سه روش، شامل روش پیشنهادی، روش اسمیت و یک حلقه پس خور معمولی مورد بررسی قرار می گیرند. معادلات (۶) تا (۸) به ترتیب مربوط به سیگنالهای ورودی به کنترلر در روش پیشنهادی، اسمیت و حلقه پس خور معمولی می باشند.

$$\varepsilon(s) = \frac{Y_{sp}(s)}{1 + G_{mb}(s) + G_c(s)G_f(s)G_0(s).e^{-sT_d}G_m(s)} \quad (6)$$

$$\varepsilon(s) = \frac{Y_{sp}}{1 + G_c(s)G_f(s)G_0(s)G_m(s)} \quad (7)$$

$$\varepsilon(s) = \frac{Y_{sp}(s)}{1 + G_c(s)G_f(s)G_0(s).e^{-sT_d}G_m(s)} \quad (8)$$

در معادلات ۶ و ۸ برای اینکه بهتر بتوان با روش اسمیت روابط را مقایسه کرد تابع فرآیند به شیوه روش اسمیت یعنی طوری که پارامتر تاخیر زمان به

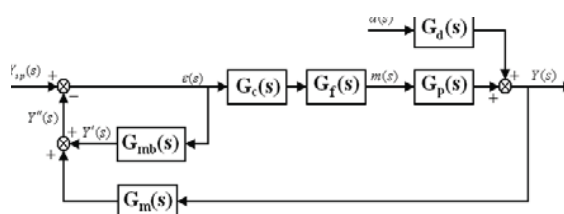
شد. برای بیان بهتر عبارتهای بالا می توان از روابط زیر استفاده نمود. اگر فرض کنیم اغتشاش ورودی به سیستم مساوی صفر است و $G_m(s) = 1$ و $G_p = G_0 e^{-sT_d}$ و سیگنال ورودی به مقایسه کننده که نهایتا وارد کنترل کننده می شود را با Y'' نمایش دهیم، داریم:

$$\begin{aligned} Y'' &= Y' + Y \\ &= m(s)G_0(s)(1 - e^{-sT_d}) \\ &\quad + m(s)G_0(s)e^{-sT_d} \\ &= m(s)G_0(s) \end{aligned} \quad (1)$$

همانطور که در (۱) دیده می شود پارامتر تاخیر زمانی از سیگنال ورودی به مقایسه کننده که نهایتا وارد کنترل کننده می شود حذف می شود. نهایتا معادله مشخصه مربوط به روش اسمیت به صورت رابطه (۲) می شود.

$$1 + G_c(s)G_f(s)G_m(s)G_0(s) = 0 \quad (2)$$

در رابطه بالا پارامتر تاخیر زمانی از معادله مشخصه حذف شده است. در حقیقت باید گفت که در روش اسمیت با حذف پارامتر تاخیر زمانی از معادله مشخصه، پاسخ سیستم را در مورد سیستمهای با تاخیر زمانی غالب (سیستمهایی که تابع حلقه باز سیستم کنترل آنها صفرهای سمت راست دارند) بهبود می دهیم. روش ارائه شده در این مقاله نیز بر اساس حذف صفراهای سمت راست تابع حلقه باز سیستم کنترل، با استفاده از مفهوم بردار برتر که در بخش اول مقاله توضیح داده شد، می باشد. در شکل ۲ ساختار سیستم کنترل پیشنهادی ترسیم شده است.



شکل ۲. بلوک دیاگرام پیشنهادی

توابع انتقال حلقه بسته این سیستم به صورت روابط (۳) و (۴) می باشند.

$$\frac{Y(s)}{Y_{sp}(s)} = \frac{G_c(s)G_f(s)G_p(s)}{1 + G_{mb}(s) + G_c(s)G_f(s)G_p(s)G_m(s)} \quad (3)$$

$$\frac{Y(s)}{d(s)} = \frac{G_d(s) + G_{mb}(s)G_d(s)}{1 + G_{mb}(s) + G_c(s)G_f(s)G_p(s)G_m(s)} \quad (4)$$

می توان با استفاده از معادله (۳) مقدار نهایی پاسخ به یک تغییر پله ای واحد در مقدار مقرر را پیش بینی کرد. اگر کنترل کننده یک کنترل کننده PI باشد این مقدار برابر با $1/K_m$ می شود، که در اینجا K_m

بسیاری در پاسخ سیستم داشته باشد که حتی در مواردی باعث ناپایداری حلقه کنترل می شود. در صورت خطا در مدل این خطا در معادله مشخصه ظاهر شده و ممکن است باعث ایجاد صفرهای سمت راست در آن شود. خطا در مدل را میتوان با معادله زیر نشان داد [۳۰].

$$P(s) = P_n(s) + \delta P(s) \quad (9)$$

در معادله (۹) $P(s)$ تابع انتقال فرآیند و $\delta P(s)$ اختلاف تابع انتقال واقعی و اسمی فرآیند می باشند. $\delta P(s)$ را به عنوان شعاع مقاوم بودن (Robustness) می شناسند و معیاری برای بررسی میزان مقاومت روش در مقابل خطای مدل می باشد [30]. با جایگذاری $P(s)$ در معادله مشخصه روش پیشنهادی، روش اسمیت و همچنین حلقه پس خور معمولی به ترتیب می توان به معادلات (۱۰) و (۱۱) و (۱۲) دست یافت.

$$\left| \delta P(s) \right|_A = \frac{|1 + G_{mb}(s) + G_c(s)P_n(s)|}{|G_c(s)|} \quad (10)$$

$$\left| \delta P(s) \right|_B = \frac{|1 + G_c(s)G_0(s)|}{|G_c(s)|} \quad (11)$$

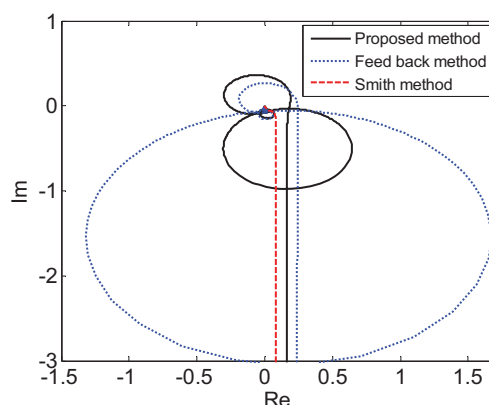
$$\left| \delta P(s) \right|_C = \frac{|1 + G_c(s)P_n(s)|}{|G_c(s)|} \quad (12)$$

با مقایسه معادلات (۱۰) و (۱۱) و (۱۲) می توان به این نتیجه رسید که در شرایط مساوی، شعاع مقاوم بودن در روش پیشنهادی نسبت به روش اسمیت و پس خور معمولی بزرگتر است. همچنین روش پیشنهادی با مقادیر بهره بزرگتر برای $G_{mb}(s)$ شعاع بزرگتری از مقاوم بودن را دارا خواهد بود. گذشته از این شعاع مقاوم بودن برای روش پیشنهادی، می تواند متغیر باشد که نشان دهنده انعطاف پذیری بالای این روش است.

۵- شبیه سازی

در این بخش با استفاده از روش پیشنهادی نمونه هایی از فرآیندهای مختلف شبیه سازی شده و در قالب این شبیه سازیها خواص روش پیشنهادی مورد بررسی قرار می گیرد. همانطور که قبلاً گفته شد در شبیه سازیها اولین کارانتخاب بهره $G_{mb}(s)$ می باشد. این بهره می بایست بزرگتر از جملات دیگر در حلقه باز سیستم کنترل باشد. البته همانطور که میدانیم هم مقدار بهره و هم ثابت زمانی یک تابع درجه یک می تواند در اندازه آن موثر باشد اما برای سادگی کار، ثابت زمانی تابع $G_{mb}(s)$ مقدار ۱ در نظر گرفته می شود. بدین ترتیب تنها مقدار بهره، در اندازه تابع موثر خواهد بود. با توجه به اینکه ما هنوز پارامترهای کنترل کننده را نداریم، ابتدا مطابق معادله (۱۳) مقدار بهره $G_{mb}(s)$ را به عنوان پیش

صورت جدا باشد نگاشته شده است. از بررسی معادلات بالا می توان به نکات جالبی در مورد پاسخ سیستم کنترل در این سه روش مختلف دست یافت. به دلیل وجود پارامتر تاخیر زمان در مخرج (۸)، ممکن است این عبارت شامل صفرهای تاخیر زمانی دار سمت راست شده و بدین ترتیب سیگنال ورودی به کنترلر حالت نامینیم فاز گرفته و پاسخ سیستم کنترل ضعیف شود. بنابراین روش اسمیت که در آن مخرج سیگنال ورودی به کنترلر فاقد صفر سمت راست است از یک حلقه پس خور معمولی بهتر عمل می کند. در روش پیشنهادی نیز به دلیل اینکه اندازه تابع G_{mb} طوری انتخاب می شود که رفتار سیستم بدون تاخیر زمانی شود، لذا وجود صفرهای سمت راست در تابع حلقه باز سیستم کنترل منتفی است. بنابراین سیگنال ورودی به کنترلر فاقد صفر سمت راست خواهد بود. برای مقایسه این سه سیگنال در شکل ۳ یک مثال عددی آورده شده است. در این شکل مدل فرآیند $G_p(s) = 4 \exp(-s)/(0.5s + 1)$ و پیش بین کننده $G_{mb}(s) = 2/(s + 1)$ میباشند.



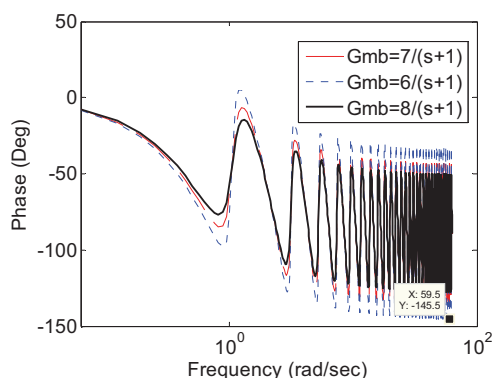
شکل ۳. مقایسه نمودار نایکیست سیگنالهای ورودی به کنترلر در حالتی، یک حلقه کنترل معمولی، پیش بین کننده اسمیت در حالتیکه در مدل خطایی وجود ندارد و روش پیشنهادی

همانطور که در این شکل دیده می شود سیگنال ورودی به کنترلر در حالت معمولی سیگنالی است که نمودار نایکیست آن نقطه -۱ را دور زده است اما در روش اسمیت و روش پیشنهادی این گونه نیست و سیستم کنترل در این دو روش مشکلی از لحاظ پایداری پیدا نخواهد کرد.

۴- مقاوم بودن روش (Robustness)

در این بخش مقاوم بودن روش نسبت به خطای مدل مورد بررسی قرار می گیرد و با دو روش مطرح شده در بالا مقایسه خواهد شد. در سیستمهای کنترل دارای تاخیر زمانی وجود خطا در مدل می تواند تاثیر

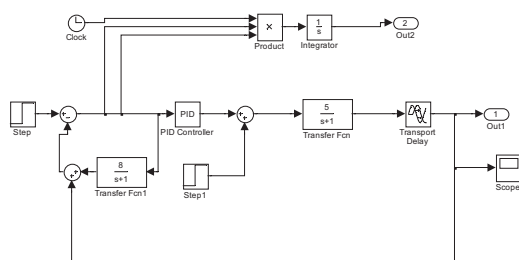
بنابراین می توان بهره های ۷ و ۸ و غیره که بزرگتر از بهره کلی حلقه باز ساده باشد نیز به عنوان بهره G_{mb} انتخاب گردد. بهر حال، از لحاظ مقاوم بودن نسبت به خطای مدل هر چه بهره بالاتر انتخاب شود پاسخ بهتر و مقاوم تر خواهد بود. بعنوان مثال نمودار فاز حلقه باز سیستم کنترل (مجموع G_p و G_{mb}) در شکل ۵ نشان داده شده است.



شکل ۵. نمودار فاز برای مقادیر مختلف بهره تابع $G_{mb}(s)$

همانطور که از شکل بالا مشخص است هر چه مقدار بهره $G_{mb}(s)$ بالا تر رود شدت نوسانات منحنی فاز کمتر شده و حاشیه فاز آن از مقدار بحرانی 180° درجه بیشتر می شود. همانطور که قبلاً نیز ذکر شد برای از بین بردن کامل صفرهای سمت راست، ما سعی می کنیم بهره $G_{mb}(s)$ را طوری تعیین کنیم که حداقل مقدار فاز آن نزدیک به 90° درجه باشد. در این مثال حداقل فاز هنگامیکه بهره ۶ است 146° - و برای بهره ۷ مقدار 135° - و بهره ۸ مقدار 127° - است بنابراین مقدار بهره ۸ را انتخاب می کنیم.

با توجه به اینکه در حال حاضر برای روش کنترل پیشنهادی هیچگونه روش خاصی جهت تنظیم و تیون کردن کنترلر وجود ندارد در این مقاله از روش انتگرال خطا برای تیون کردن پارامترهای کنترلر استفاده شده است. با استفاده از معیار ITSE و شبیه سازی توسط جعبه ابزار Simulink در Matlab-7 این شبیه سازی انجام شده است. نمودار Simulink مربوطه در شکل ۶ نمایش داده شده است.



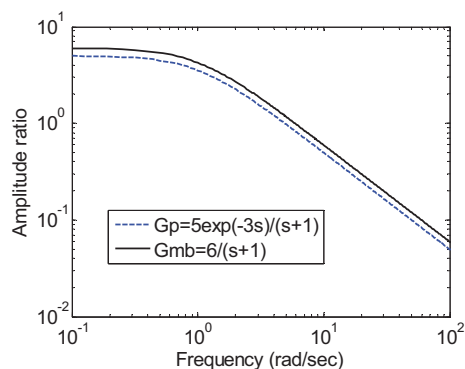
فرض اولیه به دست آورده و سپس پارامترهای کنترلر را طوری که به پاسخ مطلوب رسیده و اندازه $G_{mb}(s)$ نسبت به جملات دیگر حلقه باز سیستم کنترل برتر باشد، به دست می آوریم.

$$|G_{mb}(s)| > |G_p(s)| \quad \forall \omega \quad (13)$$

لازم به ذکر است که $G_p(s)$ نماد توابعی از سیستم کنترل است که معلوم می باشد بدیهی است که اگر در سیستم $G_f(s)$ یا $G_m(s)$ نیز موجود باشد باید حاصلضرب $G_p(s)G_f(s)G_m(s)$ را به جای $G_p(s)$ در نظر گرفت. این روش، روش دلخواهی است که در این مقاله برای پیدا کردن پارامترهای کنترلر استفاده می شود، البته باید گفت که به طرق دیگر نیز می توان پارامترهای کنترلر را پیدا کرد. مهم برتر بودن تابع $G_{mb}(s)$ نسبت به جملات دیگر تابع حلقه باز می باشد. همانطور که از روابط (۳) و (۴) مشخص است مجموع $G_p(s)$ و $G_{mb}(s)$ در اصل حلقه باز سیستم موجود در شکل ۲ می باشد لذا در این مقاله برای شبیه سازیها و پیدا کردن بهره $G_{mb}(s)$ از نمودارهای فاز و نسبت دامنه حلقه باز شکل ۲ استفاده می کنیم.

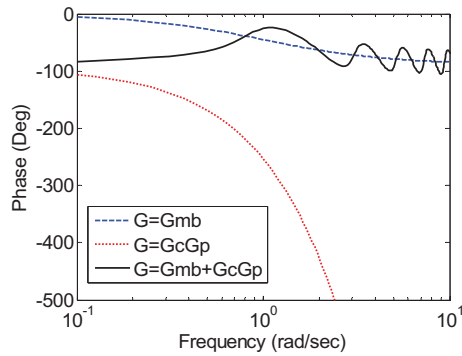
۵-۱- خاصیت مقاوم بودن روش نسبت به تغییر پارامترهای مدل

با توجه به آن که در این روش مقدار بهره تابع $G_{mb}(s)$ را باید بر اساس بهره تابع انتقال فرآیند و منحنیهای نسبت دامنه و فاز حلقه باز سیستم تعیین نمود، لذا، این مقدار را می توان طوری تعیین کرد که پاسخ سیستم نسبت به تغییرات در پارامترهای فرآیند مقاوم باشد به عنوان مثال تابع $G_p(s) = 5\exp(-3s)/(s+1)$ را در نظر می گیریم. بهره تابع $G_{mb}(s)$ پیشنهادی می تواند ۶ باشد زیرا که مطابق شکل ۴ در تمامی فرکانسها اندازه آن بالاتر از اندازه می باشد.



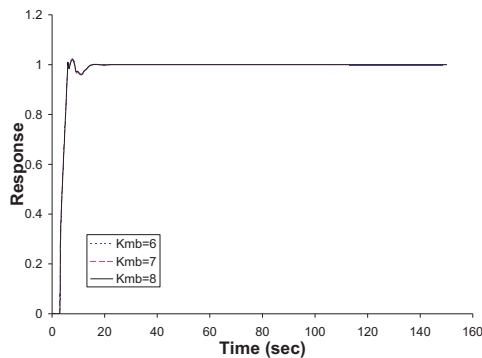
شکل ۴. نمودار نسبت دامنه برای تابع G_{mb} با بهره ۶ و تابع فرآیند

شکل ۶. بلوک دیاگرام شبیه سازی مدل با استفاده از نرم افزار سیمولینک



شکل ۷. نمودار نسبت دامنه و فاز مثال مطرح شده

در شکل ۸ پاسخ این روش با استفاده از بهره های متفاوت ترسیم شده است همانطور که مشاهده می شود پاسخها تقریباً یکسان است. در شکل ۹ پاسخ سیستم با بهره ۸ با روش اسمیت در یک نمودار ترسیم شده است. کنترلر اسمیت در اینجا با استفاده از روش Majhi, Atherton [15,18] که یک روش بهبود یافته اسمیت می باشد تون شده است و پارامتر تناسبی ۰.۲ و انتگرالی ۱ انتخاب شده است.



شکل ۸. پاسخ سیستم کنترل شده برای بهره های متفاوت ۶ و ۷ و ۸

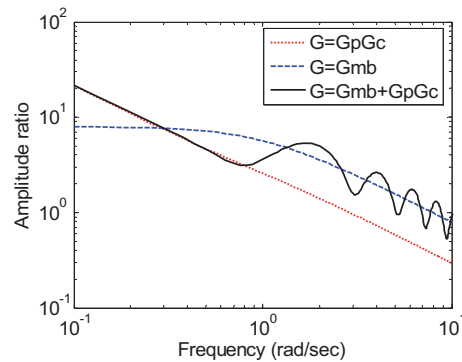
اگر در بهره مدل سیستم ۲۰٪ درصد خطا وجود داشته باشد یعنی به جای عدد ۵ این بهره ۶ باشد و در پارامتر تأخیر زمانی ۳۳٪ درصد خطا وجود داشته باشد یعنی بجای ۳ عدد ۴ باشد در این صورت روش اسمیت به طور کامل ناپایدار خواهد شد. این مطلب در شکل ۱۰ به وضوح نشان داده شده است در حالیکه روش پیشنهادی ناپایدار نشده و پس از گذشت زمانی نه چندان طولانی به مقدار مقرر می رسد. پاسخ سیستم در روش پیشنهادی برای سیستم مذکور در شکل ۱۱ نشان داده شده است.

در شکل ۶ پارامترهای کنترل کننده با استفاده از معیار ITSE بهینه می شود. در این روش پارامترهای تیونینگ کنترلر طوری ایتیم شده اند که خروجی شماره ۲ به حداقل مقدار خود برسد. در جدول ۱ پارامترهای به دست آمده برای کنترلر در K_{mb} های متفاوت نگاشته شده است.

جدول ۱. پارامترهای به دست آمده برای کنترلر در K_{mb} های متفاوت

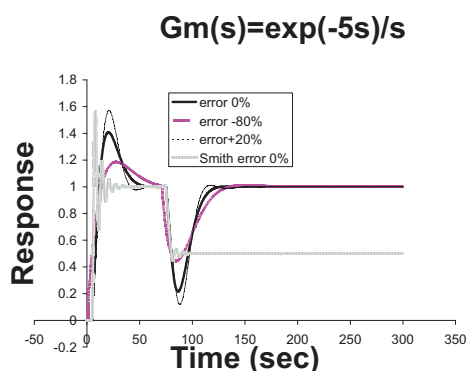
K_{mb}	پارامتر تناسبی کنترلر	پارامتر انتگرالی کنترلر
۶	۰.۴۶۱۲	۰.۳۳۳
۷	۰.۵۲۷۷	۰.۳۸۰۲
۸	۰.۵۷۹۲	۰.۴۲۷۶

مطابق جدول ۲ هرچه مقدار K_{mb} بیشتر می شود دیگر پارامترهای کنترلر نیز بالاتر می رود. در شکل ۷ نمودار نسبت دامنه و فاز تابع حلقه باز سیستم کنترل نشان داده شده است. همانطور که در نمودار نسبت دامنه دیده می شود در فرکانسهای خیلی کوچک اندازه G_{mb} از $G_p G_c$ کوچکتر است اما چون منحنی فاز نهایتاً بین -۱۸۰ و -۹۰ محدود می شود لذا هیچگونه صفر سمت راستی در تابع حلقه باز سیستم کنترل ایجاد نمی شود.



۵-۲- توانایی کنترل فرآیندهای انتگرالی با تاخیر زمانی به هنگام بروز آشفتگی

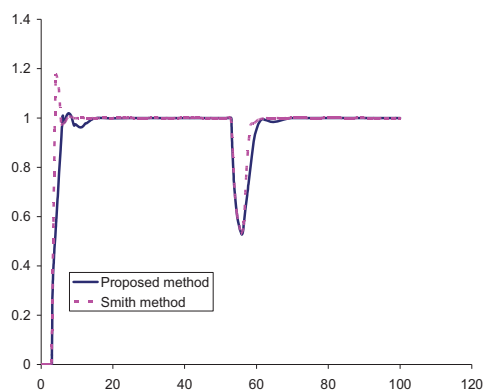
همانطور که قبلاً نیز در این مقاله ذکر شد در پاسخ سیستم کنترل شده با روش اسمیت، هنگامیکه برای فرآیندهای انتگرالی به کار می رود، افست دیده می شود اما روش پیشنهادی در این مقاله دارای این مشکل نیست. به عنوان مثال در شکل ۱۲ برای فرآیند $G_p(s) = e^{-5s}/s$ و پیش بین کننده $G_{mb}(s) = 32/(s+1)$ و کنترل کننده $G_c(s) = 4 + 0.2/s$ پاسخ سیستم کنترل شبیه سازی شده است. همانطور که در این شکل مشخص است در روش پیشنهادی بعد از ایجاد آشفتگی آفست به وجود نمی آید. همچنین مقاوم بودن روش به خطای مدل در پارامتر تاخیر زمانی نیز در این شکل نشان داده شده است. لازم به ذکر است در این سیستم کنترل، در ثانیه ۷۰ آشفتگی بمقدار ۰/۱- وارد شده است.



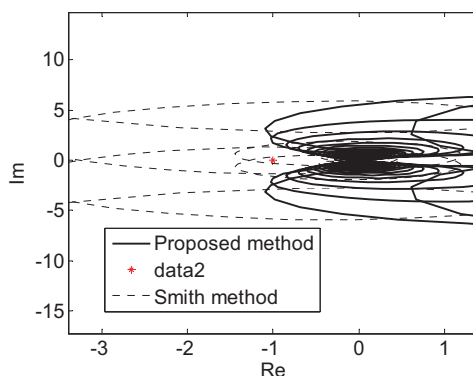
شکل ۱۲. پاسخ روش پیشنهادی برای یک فرآیند انتگرالی در هنگام بروز آشفتگی

۵-۳- قابلیت کنترل فرآیندهای پارامتر گسترده

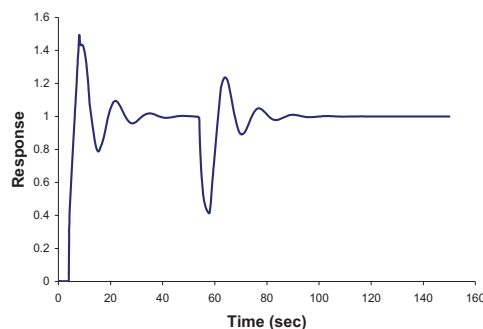
همانطور که قبلاً نیز ذکر شد این گونه سیستمها با استفاده از روش اسمیت و یا روشهای اصلاح شده اسمیت در سالهای اخیر، قابل کنترل نیستند. همچنین روشهایی که در آنها از طریق حذف صرفه‌های سمت راست تابع حلقه باز سیستم، کنترل صورت می گیرد، علاوه بر اینکه جامع نیستند مشکل و غیر دقیق نیز می باشند. در حالیکه روش پیشنهادی قادر به کنترل این سیستمها به طریقی بسیار ساده و ریاست می باشد. در شکل ۱۳ به عنوان نمونه پاسخ فرآیند $G_p = 2/(s+1) + e^{-5s}/(s+1)$ با کنترل کننده $G_c = 1 + 1/s$ نشان داده شده است. همانطور که از پاسخ مشخص است پاسخ به سرعت به مقدار مقرر باز گشته و در موارد وجود خطا بسیار ریاست می باشد. لازم به ذکر است که در دقیقه ۷۰ یک آشفتگی به مقدار



شکل ۹. مقایسه پاسخ روش اسمیت و روش پیشنهادی هنگامیکه هیچ خطایی در مدل نمی باشد.



شکل ۱۰. نمودار نایکست روش پیشنهادی و روش اسمیت در صورتیکه در پارامتر بهره مدل $+20\%$ درصد خطا و در پارامتر تاخیر زمانی $+33\%$ درصد خطا باشد داده نقطه ۱- است



شکل ۱۱. پاسخ سیستم کنترل در روش پیشنهادی در هنگام وجود خطا

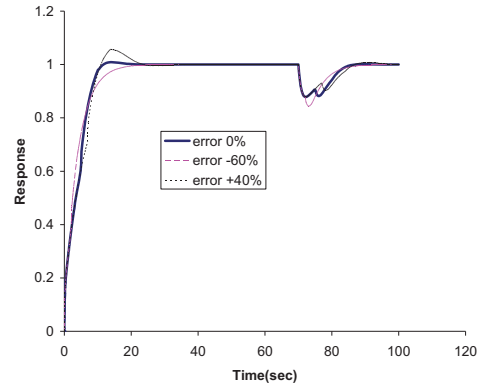
بنابراین با توجه به مثالهای مطرح شده می توان نتیجه گرفت روش پیشنهاد شده در مقایسه با روش اسمیت نسبت به خطای مدل مقاومتر بوده و حتی در مواردیکه روش اسمیت قادر به کنترل سیستم نیست روش پیشنهادی می تواند سیستم را کنترل کند.

همانطور که در شکل بالا نشان داده شده است روش نسبت به تغییرات مدل نیز مقاوم می باشد.

۰/۱- بر سیستم اعمال شده است. همانطور که از شکل نیز مشخص است باز هم آفست کاملاً بر طرف شده است.

۶- بحث و نتیجه گیری

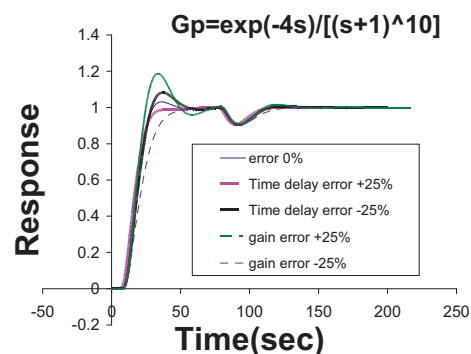
روشی که در این مقاله ارائه شده است روشی نوین در جهت کنترل فرآیندهای دارای تأخیر زمانی است. این روش که الهام گرفته از رفتار دینامیکی فرآیندهای QRDS و بر اساس انتخاب گین برتر می باشد در مقایسه با روشهای دیگر کنترل فرآیندهای دارای تأخیر زمانی مانند روش اسمیت و روشهای حذف صفرهای سمت راست از تابع حلقه بازسیستم کنترل، بسیار ساده و ریاست می باشد و به راحتی می تواند برای کنترل کلیه فرآیندهای پایدار دارای تأخیر زمانی که شامل فرآیندهای انتگرالی و پارامتر گسترده نیز می شود، در حلقه های کنترل SISO بکار رود.



شکل ۱۳. پاسخ روش پیشنهادی برای یک فرآیند پارامتر گسترده

۵-۴- توانایی رفع اثرات دینامیکی درجات بالا در مدلها

همانطور که می دانیم درجات بالا در مدل باعث ایجاد رفتاری مشابه رفتار تأخیر زمانی دار در نمودار فاز حلقه باز سیستم کنترل شود. در این مورد هم روش پیشنهادی، می تواند فاز سیستم کنترل را محدود کرده و مانع از این شود که درجه بالای مدل باعث ایجاد رفتار تأخیر زمانی دار در سیستم کنترل شود. این موضوع در شکل ۱۴ برای فرآیند $G_p(s) = e^{-4s} / [(s+1)^{10}]$ و پیش بین کننده $G_{mb}(s) = 6/(s+1)$ و کنترل کننده $G_c(s) = 2 + 0.4/s$ ، برای یک تغییر پله ای واحد در زمان صفر و آشفتگی ۰/۱- در زمان ۷۰ ثانیه، نشان داده شده است.



شکل ۱۴. کنترل مدلها با درجات بالا با استفاده از روش پیشنهادی

مراجع

- [1] Richard J.- P., "Time delay systems: an overview of some recent advances and open problems", Automatica, Vol. 39, pp.1667-1694, 2003.
- [2] Normey-Rico J.E. & Camacho F., Eduardo, "Dead-time compensators: A survey." Control Engineering Practice, doi:10.1016/j.conengprac.2007.05.006.
- [3] Smith, O. J. M. "Closer control of loops with dead time", Chemical Engineering Progress, Vol. 53, 1959, pp.217-219.
- [4] Palmor, Z., "Stability Properties of Smith Dead-Time Compensator Controllers," Int. J. Control, Vol. 32, No. 6, 1980, p. 212.
- [5] Normey-Rico, J. E., Bordons, C., & Camacho, E.F., "Improving the robustness of dead time compensating PI controllers" Control Engineering Practice, Vol. 5, No. 6, 1997, pp. 801-810.
- [6] Ingimundarson, A., & Häggglund, T., "Robust tuning procedure of dead time compensating controllers", Control Engineering Practice, No 9, 2001, pp. 1195-1208.
- [7] Michiels, W., & Niculescu, S.i. "On the delay sensitivity of smith predictors", International Journal of Systems Science, Vol 34(8-9), 2003, pp. 543-552.
- [8] Watanabe, K., and M. Ito, "A Process - Model Control for Linear Systems with Time-Delay," IEEE Trans. Automat. Contr., Vol. AC - 26, No. 6, Dec. 1981, pp.1261 - 1269.
- [9] Astrom, K., Hang, C. C., & Lim, B. C.. "A New Smith Predictor for Controlling A Process with an Integrator and Long Dead-Time," IEEE Transactions

- Distributed Parameter Systems," *AICHe J.*, Vol. 41, 1995, pp. 2658-2660.
- [24] Shirvani, M., M. A. Doustary, M. Shahbaz and Z. Eksiri, "Heuristic Process Model Simplification In Frequency Response Domain," *I. J. E. Transactions B: Applications*, Vol. 17, No. 1, 2004, pp.19-39.
- [25] Vollmer U, Raisch J., "H_∞-Control of a continuous crystallizer", *Control Engineering Practice*, 9, , 2001, pp.837-845.
- [26] Zltek P., Hlava J., "Anisochronic internal model control of time delay systems ", *Control Engineering Practice*, No. 9, 2001, pp.501-516.
- [27] Vollmer U., Raisch J., "Population balance modeling and H_∞-Controller design for a crystallization process", *Chemical Engineering Science*, No.57, 2002, pp.4401-4414.
- [28] Dimitri Breda," Solution operator approximations for characteristic roots of delay differential equations", *Applied Numerical Mathematical*, 56, 2006, pp.305-317.
- [29] K. Verheyden, T. Luzyanina and D. Roose," Efficient computation of characteristic roots of delay differential equations using LMS methods, *J. Comput. Appl. Math.*, 2007, doi:10.1016/j.cam.2007.02.025.
- [30] Morari, M., & Zafriou, E." Robust process control", Englewood Cliffs, NJ: Prentice-Hall.
- on Automatic Control, Vol 39, No 2, 1994, pp. 343-345.
- [10] Ingimundarson, A., & Hägglund, T. " Performance comparison between PID and dead time compensating controllers", *Journal of Process Control*, No.12, 2002,pp. 887-895.
- [11] Hägglund, T., "A predictive PI Controller for Processes with Long Dead-Times," *IEEE Contr. Syst. Mag.*, Vol. 12, No.1, 1992, pp. 57 - 60.
- [12] Kaya, I. and D. P. Atherton, "Anew PI - PD Smith Predictor for Control of Processes with Long Dead-Time", in: 14th IFAC World Congress, Vol. C, 1999, PP.283-288.
- [13] Kaya, I., "Obtaining Controller Parameters for a New PI-PD Smith Predictor Using Auto tuning," *Journal of Process Control*, No, 13, 2003 pp. 465-472
- [14] Majhi, S. and D. P. Atherton, "Modified Smith Predictor and Controller for Processes with Time-Delay," *IEE Proc. Control Theory Appl.* Vol.146, No. 5, 1999, pp.359-366.
- [15] Majhi, S. and D. P. Atherton, "Obtaining Controller Parameters for a New Smith Predictor Using Auto Tuning," *Automatica* No.36, 2000, pp.1651-1658.
- [16] Mataušek, M. R. and A. D. Micić, "A modified Smith Predictor for Controlling a Process with an Integrator and Long Dead-Time," *IEEE Trans. Automat. Control*, No. 44, 1996, pp. 1196-1203.
- [17] Mataušek, M . R. and A. D. Micić, "On the Modified Smith Predictor for Controlling a Process with an Integrator and Long Dead-Time," *IEEE Trans. Automat. Cont.*, 41, 1999, pp.1603-1606.
- [18] Xiang , Lu., Yang,Y., Wang,Q and Zheng, W., "A Double Two - Degree - of -Freedom Control Scheme for Improved Control of Unstable Delay Processes," *Journal of Process Control*, No. 15, 2005, pp. 605 - 614.
- [20] Ramanathan, S., "Control of Quasirational Distributed Systems with Examples on the Control of Cumulative Mass Fraction of Particle Size Distribution,"Ph.D. Thesis, University of Michigan, Ann Arbor, 1988.
- [21] Ramanathan, S., Curl R. L. and C. Kravaris "Dynamics and Control of Quasirational Systems," *AICHe J.*, Vol. 55, 1989.
- [22] Shirvani, M., M. Inagaki and T. Shimizu, "A Simplified Model of Distributed Parameter Systems," *Int. J. Eng.*, Vol. 6, No. 2, 1993.
- [23] Shirvani, M, M. Inagaki, and T. Shimizu, "Simplification Study on Dynamic Models of

New Steady State Kalman Filter for Tracking High Maneuvering Targets

Seyyed-Amin Hosseini¹, Mohammad Farrokhi^{1,2}

¹ Department of Electrical Engineering

² Center of Excellence for Power System Operation and automation

Iran University of Science and Technology, Tehran 16846, Iran,

amin_hosseini@iust.ac.ir, farrokhi@iust.ac.ir

Abstract: This paper presents a new steady state Kaman filter for tracking high maneuvering targets. The $\alpha - \beta$ and $\alpha - \beta - \gamma$ filters are the steady-state Kaman filters for tracking constant speed and constant acceleration targets, respectively. However, these filters cannot predict high maneuvering targets with good accuracy. The proposed filter is called $\alpha - \beta - \gamma - \eta$ filter. In this filter η is the gain of jerk for tracking of high maneuvering targets. Simulation results show good performance of the proposed filter as compared to $\alpha - \beta - \gamma$ filter to track the jerky movements.

Keywords: Target Tracking, Steady-State Estimation Filters, $\alpha - \beta$ Filter, $\alpha - \beta - \gamma$ Filter.

چکیده: در این مقاله، فیلتر جدیدی از نوع کالمن با حالت ماندگار برای ردیابی اهداف با مانور سریع ارائه می‌شود. فیلترهای $\alpha - \beta$ و $\alpha - \beta - \gamma$ فیلترهای از نوع کالمن با حالت ماندگار به ترتیب برای ردیابی اهداف با سرعت ثابت و شتاب ثابت هستند. لیکن، این فیلترها قادر به ردیابی اهداف دارای مانور سریع با دقت خوب نیستند. فیلتر پیشنهادی در این مقاله فیلتر $\alpha - \beta - \gamma - \eta$ نامیده می‌شود که در آن η بهره جریک برای ردیابی هدف می‌باشد. نتایج شبیه‌سازی، عملکرد خوب فیلتر پیشنهادی را در مقایسه با فیلتر $\alpha - \beta - \gamma$ در ردیابی مانورهای جریکی نشان می‌دهد.

کلمات کلیدی: ردیابی هدف، فیلترهای تخمین‌زن با حالت ماندگار، فیلتر $\alpha - \beta$ ، فیلتر $\alpha - \beta - \gamma$

Nomenclatures:

\mathbf{x} : state vector ($n \times 1$)
 \mathbf{F} : transformation matrix ($n \times n$)
 \mathbf{u} : process noise
 \mathbf{v} : measurement noise
 x : position state
 \dot{x} : velocity state
 \ddot{x} : acceleration state
 \dddot{x} : jerk state
 T : sampling time
 \mathbf{h} : observability matrix ($3 \times n$)
 a : correlation parameter
 Q_j : variance of process noise
 \mathbf{Q} : covariance matrix of process noise ($n \times n$)
 N : variance of target jerk
 σ_j^2 : variance of target jerk
 R : variance of measurement noise
 σ_v^2 : variance of measurement noise
 \mathbf{P} : state covariance matrix ($n \times n$)
 P_{ij} : elements of \mathbf{P}

\mathbf{S} : updated state covariance matrix ($n \times n$)
 \mathbf{w} : gain of filter
 g_i : elements of \mathbf{w}
 α : position gain
 β : velocity gain
 γ : acceleration gain
 η : jerk gain
 λ : tracking index

1- Introduction

The key point to successful target tracking is to extract useful information about the target state from the observed data. In order to achieve this goal, one needs a useful target model. The simplest model for a target, the so-called *white-noisy acceleration model* is used when the maneuver is small or random [1], [2]. The other simple model is the *Wiener-process acceleration model* that is referred to as the *constant-acceleration model* [1], [2]. On the other hand, the Singer acceleration model is a standard model for targets with maneuvers [3]. This model assumes that the target acceleration is a zero-mean stationary first-

order Markov process. For high maneuvering targets, a jerky model is proposed by Mehrotra and Mahapatra [4]. In this method, the jerk is modeled as a zero-mean first-order Markov process, in the same way as the Singer acceleration model.

One of the widely used approaches for state estimation is the Kalman filter [5], [6]. However, Kalman filter imposes large amount of computations. In order to reduce computational burdens, constant-gains Kalman filters are used. The $\alpha-\beta$ and the $\alpha-\beta-\gamma$ filters are the steady-state Kaman filters for target tracking with constant speeds and constant accelerations, respectively [7]-[10]. For high maneuvering targets (i.e. targets with changing accelerations), the performance of these steady-state filters deteriorates. To improve this shortcoming, an extension of these filters is proposed in this paper. This filter is called the $\alpha-\beta-\gamma-\eta$ filter, in which η is the gain for jerk. Simulation results show good performance of the proposed filter as compared to the $\alpha-\beta-\gamma$ filter for tracking jerky movements.

This paper is organized as follows. Section 2 defines the jerk model. Section 3 presents gain computations for the proposed filter. Section 4 gives the simulation results. Finally, Section 5 concludes the paper.

2- Jerk Model for Target Motion

There are many types of target motion, especially those involving the modern generation of highly maneuvering aerospace vehicles that call for better tracking performance than what is provided by acceleration models. The reason for the inadequate tracking performance of current models is that the higher order derivatives in the case of very highly maneuvering targets are not insignificant, leading to model inaccuracies when terms only up to the second or third orders derivatives are considered.

The state model for the target motion is defined by the following vector-matrix equations [4]:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{F}(k)\mathbf{x}(k) + \mathbf{u}(k), \quad (1) \\ y(k+1) &= \mathbf{h}(k+1)\mathbf{x}(k+1) + v(k+1), \quad (2) \end{aligned}$$

where

$\mathbf{x} = [x \quad \dot{x} \quad \ddot{x} \quad \ddot{\ddot{x}} \quad y \quad \dot{y} \quad \ddot{y} \quad \ddot{\ddot{y}} \quad z \quad \dot{z} \quad \ddot{z} \quad \ddot{\ddot{z}}]^T$ is the state vector, y is the output, \mathbf{u} and v are white noises denoting the process and the measurement noises, respectively, all with appropriate dimensions. Moreover, the transition, observation, and covariance matrices are

$$\mathbf{F} = \begin{bmatrix} 1 & T & \frac{T^2}{2} & \frac{T^3}{6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & T & \frac{T^2}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & T & \frac{T^2}{2} & \frac{T^3}{6} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & T & \frac{T^2}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & T & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & T & \frac{T^2}{2} & \frac{T^3}{6} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & T & \frac{T^2}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & T & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (3)$$

$$\mathbf{h} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad (4)$$

$$\mathbf{Q} = \mathbf{Q}_j \begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ q_{12} & q_{13} & q_{14} & q_{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ q_{13} & q_{14} & q_{15} & q_{16} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ q_{14} & q_{15} & q_{16} & q_{17} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q_{11} & q_{12} & q_{13} & q_{14} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q_{12} & q_{13} & q_{14} & q_{15} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q_{13} & q_{14} & q_{15} & q_{16} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & q_{14} & q_{15} & q_{16} & q_{17} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{11} & q_{12} & q_{13} & q_{14} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{12} & q_{13} & q_{14} & q_{15} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{13} & q_{14} & q_{15} & q_{16} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q_{14} & q_{15} & q_{16} & q_{17} & 0 \end{bmatrix}, \quad (5)$$

where $q_{11} = T^7/252$, $q_{12} = T^6/72$, $q_{13} = T^5/30$, $q_{14} = T^4/24$, $q_{15} = T^3/6$, $q_{16} = T^2/2$, $q_{17} = T$. Moreover, T is the sampling time interval and $\mathbf{Q}_j = 2a\sigma_j^2$ is the variance of the process noise in the jerk model, in which a is the correlation parameter and σ_j^2 is the variance of the target jerk. Matrix \mathbf{h} in (4) shows that only the position sensor is used for target tracking.

In the next section, for the ease of presentation, only the state variables along the x axis (i.e. x , \dot{x} , \ddot{x} , and $\ddot{\ddot{x}}$) will be considered. For other axes (i.e. y and z) the same equations apply. This fact can also be observed from matrices \mathbf{F} and \mathbf{Q} in (3) and (5), respectively.

3- The $\alpha-\beta-\gamma-\eta$ Filter

The following equations hold for the Kalman filter in steady-state conditions [9]:

$$\mathbf{P}(k|k) = \mathbf{P}(k-1|k-1),$$

$$\mathbf{P}(k+1|k) = \mathbf{P}(k|k-1),$$

$$\mathbf{w}(k) = \mathbf{w}(k-1),$$

where \mathbf{w} is the gain of the filter. The components of the steady-state covariance matrix is denoted as

$$\lim_{k \rightarrow \infty} \mathbf{P}(k | k) = [\mathbf{P}_{ij}]. \quad (6)$$

Components of the one-step prediction of the covariance matrix are defined as

$$\lim_{k \rightarrow \infty} \mathbf{P}(k+1 | k) = [\mathbf{M}_{ij}]. \quad (7)$$

The updated covariance in Kalman filter is defined as

$$S = \mathbf{hP}(k+1 | k)\mathbf{h}^T + R, \quad (8)$$

where \mathbf{h} is the same as before, $R = \sigma_v^2$, in which σ_v^2 is the variance of the measurement noise. Using (4), (7) and (8), the updated covariance will be

$$S = m_{11} + R. \quad (9)$$

Hence, the Kalman filter gain becomes

$$\begin{aligned} \mathbf{w} &= \mathbf{P}(k+1 | k)\mathbf{h}^T S^{-1} \\ &= \begin{bmatrix} \frac{m_{11}}{m_{11} + R} & \frac{m_{12}}{m_{11} + R} & \frac{m_{13}}{m_{11} + R} & \frac{m_{14}}{m_{11} + R} \end{bmatrix}^T \\ &= [g_1 \quad g_2 \quad g_3 \quad g_4]^T. \end{aligned} \quad (10)$$

Using (10), it yields

$$m_{1i} = \frac{g_i}{1 - g_1} R \quad i = 1, 2, 3, 4. \quad (11)$$

Hence, the updated covariance matrix becomes

$$\mathbf{P}(k+1 | k+1) = (\mathbf{I} - \mathbf{wh})\mathbf{P}(k+1 | k). \quad (12)$$

Using (6), (7) and (12), it gives

$$[\mathbf{P}_{ij}] = (\mathbf{I} - \mathbf{wh})[\mathbf{M}_{ij}]. \quad (13)$$

Therefore, $[\mathbf{P}_{ij}]$ is

$$[\mathbf{P}_{ij}] = \begin{bmatrix} (1-g_1)m_{11} & (1-g_1)m_{12} & (1-g_1)m_{13} & (1-g_1)m_{14} \\ (1-g_1)m_{12} & m_{22} - g_2 m_{12} & m_{23} - g_2 m_{13} & m_{24} - g_2 m_{14} \\ (1-g_1)m_{13} & m_{23} - g_2 m_{13} & m_{33} - g_3 m_{13} & m_{34} - g_3 m_{14} \\ (1-g_1)m_{14} & m_{24} - g_2 m_{14} & m_{34} - g_3 m_{14} & m_{44} - g_4 m_{14} \end{bmatrix}. \quad (14)$$

Hence, the covariance prediction equation in Kalman filter is

$$\mathbf{P}(k | k) = \mathbf{F}^{-1} [\mathbf{P}(k+1 | k) - \mathbf{Q}] (\mathbf{F}^{-1})^T = [\mathbf{P}_{ij}]. \quad (15)$$

If $[\mathbf{P}_{ij}] = \mathbf{P}$, then, according to (14) the elements of \mathbf{P} are

$$\begin{aligned} P(1,1) &= m_{11} - \frac{1}{126} a \sigma_j^2 T^7 - 2T m_{12} + T^2 m_{13} \\ &\quad - \frac{1}{3} T^3 m_{14} + T^2 m_{22} - T^3 m_{23} + \frac{1}{3} T^4 m_{24} \\ &\quad + \frac{1}{4} T^4 m_{33} - \frac{1}{6} T^5 m_{34} + \frac{1}{36} T^6 m_{44}, \end{aligned} \quad (16)$$

$$\begin{aligned} P(1,2) = P(2,1) &= m_{12} + \frac{1}{36} aNT^6 - Tm_{22} + \frac{3}{2} T^2 m_{23} \\ &\quad - \frac{2}{3} T^3 m_{24} - Tm_{13} - \frac{1}{2} T^3 m_{33} + \frac{5}{12} T^4 m_{34} \\ &\quad + \frac{1}{2} T^2 m_{14} - \frac{1}{12} T^5 m_{44}, \end{aligned} \quad (17)$$

$$\begin{aligned} P(1,3) = P(3,1) &= m_{13} - \frac{1}{15} aNT^5 - Tm_{23} + \frac{1}{2} T^2 m_{33} \\ &\quad - \frac{2}{3} T^3 m_{34} - Tm_{14} + T^2 m_{24} + \frac{1}{6} T^4 m_{44}, \end{aligned} \quad (18)$$

$$\begin{aligned} P(1,4) = P(4,1) &= m_{14} + \frac{1}{12} aNT^4 - Tm_{33} + \frac{1}{2} T^2 m_{34} \\ &\quad - \frac{1}{6} T^3 m_{44}, \end{aligned} \quad (19)$$

$$\begin{aligned} P(2,2) &= m_{22} - \frac{1}{10} aNT^5 - 2Tm_{23} + T^2 m_{24} + T^2 m_{33} \\ &\quad - T^3 m_{34} + \frac{1}{4} m_{44}, \end{aligned} \quad (20)$$

$$P(4,4) = m_{44} - 2aNT, \quad (21)$$

$$\begin{aligned} P(2,3) = P(3,2) &= m_{23} + \frac{1}{4} aNT^4 - Tm_{33} \\ &\quad + \frac{3}{2} T^2 m_{34} - Tm_{24} + \frac{1}{2} T^3 m_{44}, \end{aligned} \quad (22)$$

$$P(3,3) = m_{33} - \frac{2}{3} aNT^3 - 2Tm_{34} + T^2 m_{44}, \quad (23)$$

$$P(3,4) = m_{34} + aNT^2 - Tm_{44}. \quad (24)$$

where $N = \sigma_j^2$ and σ_j^2 was defined before.

On the other hand, according to (14), elements of $[\mathbf{P}_{ij}]$, which are designated with $PP(i, j)$, ($i = j = 1, \dots, 4$) for convenience, are equal to

$$PP(1, l) = PP(l, 1) = m_{1l} - m_{1l} g_1, \quad l = 1, 2, 3, 4, \quad (25)$$

$$PP(2, j) = PP(j, 2) = m_{2j} - m_{2j} g_2, \quad j = 2, 3, 4, \quad (26)$$

$$PP(1, k) = PP(k, 1) = m_{3k} - m_{1k} g_3, \quad k = 3, 4, \quad (27)$$

$$PP(4,4) = m_{44} - m_{14}g_4. \quad (28)$$

Equating $P(i,j)$ and $PP(i,j)$ yields

$$m_{44} = \frac{1}{T} \frac{g_3 g_4}{1-g_1} R + aNT, \quad (29)$$

$$m_{34} = \frac{1}{2T} \frac{g_3^2}{1-g_1} R + \frac{1}{6} aNT^2 + \frac{1}{2} \frac{g_4 g_3}{1-g_1} R, \quad (30)$$

$$m_{34} = \frac{1}{T} \frac{g_2 g_4}{1-g_1} R + \frac{1}{6} aNT^2 + \frac{1}{2} \frac{g_4 g_3}{1-g_1} R, \quad (31)$$

$$m_{24} = -\frac{1}{12T} R \frac{12g_1 g_4 + 3Tg_3^2 + T^2 g_3 g_4}{(g_1-1)}, \quad (32)$$

$$m_{33} = -\frac{1}{6T} R \frac{6g_3 g_2 + 3Tg_3^2 + T^2 g_3 g_4 - 6g_1 g_4}{(g_1-1)}, \quad (33)$$

$$m_{23} = -\frac{1}{90T} (90Rg_1 g_3 - aNT^5 + aNT^5 g_1 + 45TRg_2 g_3 + 15T^2 Rg_3^2 + 45TRg_1 g_4 - 90TRg_4)/(g_1-1), \quad (34)$$

$$m_{22} = -\frac{1}{360T} (-360TRg_3 - aNT^6 - 360T^2 Rg_4 + 360g_1 g_2 R + 15T^3 g_3^2 R + 90T^2 Rg_2 g_3 + 210T^2 Rg_1 g_2 + 540TRg_1 g_3 + aNT^6 g_1 - 5T^4 Rg_3 g_4)/(g_1-1). \quad (35)$$

It should be noted that two different m_{34} are obtained in (30) and (31), which is due to the different elements of $P(i,j)$ and $PP(i,j)$. Equating (16) and (25) for $l=1$ and using (29) to (35) gives

$$-aNT^7 + aNT^7 g_1 - 1260g_1 g_3 T^2 - 420T^3 Rg_1 g_4 - 2520g_1 g_2 R + 5040Tg_2 R + 840T^3 g_4 R - 2520g_1^2 R = 0. \quad (36)$$

The gains of this filter must cancel out the effects of T in (36). The gains are selected as [1]

$$\mathbf{w} = \begin{bmatrix} \alpha & \beta & \gamma \\ T & T & T \end{bmatrix} = [g_1 \ g_2 \ g_3] \quad (37)$$

Then, using (30) and (31) g_4 can be calculated as

$$g_4 = \frac{g_3^2}{2g_2}. \quad (38)$$

Therefore, according to (37) and (38), g_4 is equal to

$$g_4 = \frac{\eta}{8T^3}, \quad (39)$$

where

$$\eta = \frac{\gamma^2}{\beta}. \quad (40)$$

In steady-state tracking filters, gains are computed based on the tracking index. To compute the tracking index and the gains, some additional equations are needed. By equating (21) and (28), another form of m_{14} can be obtained as

$$m_{14} = \frac{2aNT}{g_4}. \quad (41)$$

Using (11) and (41), a new set of equations for m_{ij} are obtained as follows:

$$m_{44} = \frac{2aN}{g_4} g_3 + aNT \quad (42)$$

$$m_{34} = \frac{1}{2T} \frac{Rg_3^2}{1-g_1} + \frac{1}{6} aNT^2 + aNT \frac{g_3}{g_4} \quad (43)$$

$$m_{34} = 2aN \frac{g_2}{g_4} + \frac{1}{6} aNT^2 + aNT \frac{g_3}{g_4} \quad (44)$$

$$m_{24} = 2aN \frac{g_1}{g_4} + \frac{1}{4} \frac{g_3^2}{1-g_1} + \frac{1}{6} aNT^2 \frac{g_3}{g_4} \quad (45)$$

$$m_{33} = \frac{1}{T} \frac{Rg_3 g_2}{1-g_1} + \frac{1}{2} \frac{Rg_3^2}{1-g_1} + \frac{1}{3} aNT^2 \frac{g_3}{g_4} - 2aN \frac{g_1}{g_4} \quad (46)$$

$$m_{23} = \frac{1}{2T} R \frac{g_2^2}{1-g_1} - \frac{1}{120} aNT^4 + \frac{1}{8} TR \frac{g_3^2}{1-g_1} + \frac{1}{2} R \frac{g_2 g_3}{1-g_1} \quad (47)$$

$$m_{22} = \frac{1}{T} \frac{Rg_1 g_2}{1-g_1} + \frac{1}{720} aNT^5 + \frac{3}{4} R \frac{g_2^2}{1-g_1} - \frac{1}{48} RT^2 \frac{g_3^2}{1-g_1} + \frac{1}{4} RT \frac{g_2 g_3}{1-g_1} - \frac{1}{3} aNT^2 \frac{g_1}{g_4} - \frac{1}{36} aNT^4 \frac{g_3}{g_4} \quad (48)$$

Equating (16) and (25) for $l=1$ and using (41) to (48) gives

$$\frac{Tg_1 g_2 R}{1-g_1} - \frac{1}{6} \frac{T^4 aNg_1}{g_4} + \frac{1}{3} \frac{T^4 aN}{g_4} + \frac{aNT^7}{560} - \frac{2TRg_2}{1-g_1} + \frac{1}{4} \frac{T^2 Rg_2^2}{1-g_1} + \frac{1}{48} \frac{T^4 Rg_3^2}{1-g_1} + \frac{Rg_1^2}{1-g_1} = 0. \quad (49)$$

To determine the tracking index, equating (42) and (12) for $i=4$ results in

$$\lambda^2 = \frac{aNT^7}{R} = \frac{\eta^2}{128(1-\alpha)}. \quad (50)$$

Using the obtained gains and the tracking index in (50) and replacing them into (36) and (49) yields

$$-\frac{1}{128}\eta^2 - 630\alpha\gamma - \frac{105}{2}\alpha\eta - 2520\alpha\beta + 5040\beta + 105\eta - 2520\alpha^2 = 0 \quad (51)$$

$$\frac{1}{71680}\eta^2 + \frac{1}{48}\eta + \alpha\beta - \frac{1}{96}\eta\alpha + \alpha^2 - 2\beta + \frac{1}{4}\beta^2 - \frac{1}{192}\gamma^2 = 0 \quad (52)$$

Now, using (40), (50), (51) and (52), it gives

$$\begin{aligned} & -16 + 16\alpha + \frac{164}{105}\lambda^4 - \frac{1544}{315}\lambda^4\alpha + \frac{596}{105}\lambda^4\alpha^2 \\ & + \frac{8}{3}\lambda^2\alpha^2 - \frac{304}{105}\lambda^4\alpha^3 + \frac{16}{105}\lambda^6 - 4\lambda^2\alpha^3 \\ & - \frac{8}{15}\lambda^6\alpha^2 + \frac{24}{35}\lambda^6\alpha^2 - \frac{8}{21}\lambda^6\alpha^3 - \frac{16}{1225}\lambda^8\alpha \\ & + \frac{24}{1225}\lambda^8\alpha^2 - \frac{16}{1225}\lambda^8\alpha^3 + \frac{4}{1225}\lambda^8\alpha^4 \\ & + \frac{8}{105}\lambda^6\alpha^4 + \frac{179}{315}\lambda^4\alpha^4 + \frac{4}{3}\lambda^2\alpha^4 + \frac{4}{1225}\lambda^8 \\ & + \alpha^4 = 0 \end{aligned} \quad (53)$$

By calculating λ (using $\lambda^2 = aNT^7/R$) and putting that into (53), α can be obtained. Then, η can be calculated using (50). Finally, β and γ can be calculated using the following equations:

$$\beta = -(2\alpha - 4 - \frac{1}{192}\eta) - \frac{1}{2} \left[(4\alpha - 8 - \frac{1}{48}\eta)^2 - 4 \left(\frac{1}{17920}\eta^2 + \frac{1}{12}\eta - \frac{1}{24}\eta\alpha + 4\alpha^2 \right) \right]^{1/2} \quad (54)$$

$$\gamma = \sqrt{\eta\beta} \quad (55)$$

It should be noted that, in the process of finding these gains, the equations usually yield multiple answers, from which only the positive answers are acceptable. Moreover, from the predefined range of each gain, one can find the correct answer from the multiple positive solutions.

4- Simulation Results

In this section, a comparison between the proposed filter and the $\alpha-\beta-\gamma$ filter is demonstrated throughout simulations. The variance of the process noise for the jerk model is $Q_j = 2a\sigma_j^2$ where $\sigma_j = 0.1m/s^3$. It is assumed that the radar measurement sequences are transformed from the polar coordinates to the Cartesian coordinates before the track-while-scan (TWS) process takes place [1]. The measurement covariance matrix is defined as

$$RR = \begin{bmatrix} R_{11} & 0 & 0 \\ 0 & R_{22} & 0 \\ 0 & 0 & R_{33} \end{bmatrix} \quad (56)$$

where R_{11} , R_{22} and R_{33} are variances in x , y , and z directions, respectively, and are equal to $2500 m^2$.

Moreover, the correlation parameter is $a = 0.7$ and the sampling time interval is $T = 0.5$ sec., which is the time of the radar antenna scanning a revolution. The jerk model, which is used for all filters, is the same model as Mehrotra and Mahapatra have proposed in [4]. For the initial position equal to $[0 \ 0 \ 1000]^T m$, and the initial velocity of $[-1000 \ 10 \ 0]^T m/sec$, the gains are calculated as

$$\alpha = 0.9493, \beta = 1.2004$$

$$\gamma = 0.0213, \eta = 3.7695e-004$$

It should be noted that the gains of the $\alpha-\beta-\gamma$ filter are not necessarily the same as the proposed $\alpha-\beta-\gamma-\eta$ filter (although the derivation procedure is similar). The reader may refer to reference [1] on how to find the gains of the $\alpha-\beta-\gamma$ filter.

The jerk of the target for different time intervals is summarized in Table 1. Since the most important issue in target tracking applications is the position error, only position errors are shown in simulation results.

Figs. 1-3 show the position errors in the x , y , and z directions, respectively. As these figures show, when the jerk is zero (i.e. the target is moving with constant acceleration) the $\alpha-\beta-\gamma$ filter performs slightly better than the proposed filter in this paper. On the other hand, when the target has jerky movements, the error of $\alpha-\beta-\gamma$ filter increases or even may loose the target. The reference tracking path for the $\alpha-\beta-\gamma-\eta$ filter is depicted in Fig. 6 in three dimensions. It should be mentioned that the same tracking path is used for the $\alpha-\beta-\gamma$ filter.

For comparison, the Root-Mean-Square (RMS) errors for 500 sec. of simulations are given in Table 2.

5- Conclusion

In this paper, a new filter, called the $\alpha-\beta-\gamma-\eta$ filter, was introduced. The proposed filter is an extension of the $\alpha-\beta-\gamma$ filter, where η is the gain for the target jerk. These filters are constant-gain filters and have lower calculation volume as

compared to the Kalman filters. In addition, their tracking accuracy is acceptable. It was shown by simulations that the proposed filter can follow jerky models with high maneuvering properties, with good accuracy as compared to the $\alpha-\beta-\gamma$ filter.

Table 1: Jerk of the target for different time intervals.

time interval		target movement
From	To	
$t = 0$ sec.	$t = 250$ sec.	Jerk = 0 m/s^3
$t = 251$ sec.	$t = 500$ sec.	Jerk = 0.25 m/s^3

Table2: RMSE on x, y and z axes in 500 sec

	RMSE on x axis	RMSE on y axis	RMSE on z axis
Proposed	50.18	42.56	43
$\alpha-\beta-\gamma$	123.3675	141	202

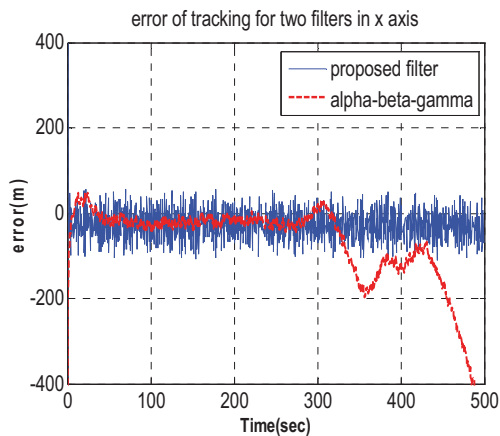


Fig. 1: Position tracking error on the x-axis

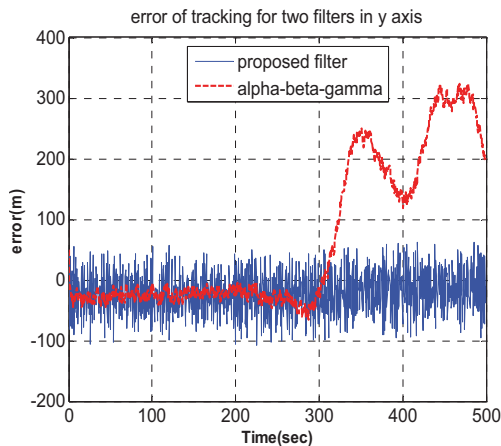


Fig. 2: Position tracking error on the y-axis

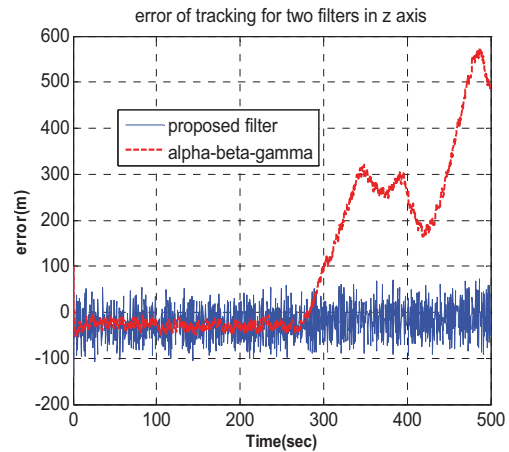


Fig. 3: Position tracking error on the z-axis

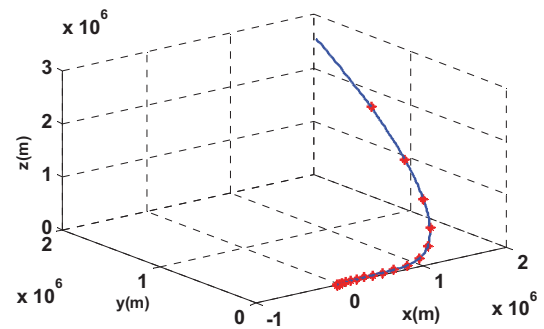


Fig.4: The model path and the tracking path in 3D

References

- [1] Bar-Shalom Y., Li X. R., and Kirubarajan T., Estimation with Applications to Tracking and Navigation, John Wiley & Sons Inc., 2003.
- [2] Rong Li X. and Jilkov V. P., "Survey of maneuvering target tracking: Part I: Dynamic models," IEEE Transactions on Aerospace and Electronic Systems, 2003, vol. 39, issue 4, pp. 1333-1364.
- [3] Singer R. A., "Estimation optimal tracking filter performance for manned maneuvering targets," IEEE Transactions on Aerospace and Electronic Systems, 1970, vol. AES-6, pp. 473-483.
- [4] Mehrotra K. and Mahapatra P. R., "A jerk model for tracking highly maneuvering targets," IEEE Transactions on Aerospace and Electronic Systems, 1997, vol. 33, issue 4, pp. 1094-1105.
- [5] Funk N., A Study of the Kalman Filter Applied to Visual Tracking, 2003, University of Alberta, Project CMPUT 652.

- [6] Hawkins M., High Speed Target Tracking Using Kalman Filter and Partial Window Imaging, 2002, Master Thesis, Georgia institute of Technology.
- [7] Arcasoy C. C., "The $\alpha - \beta - \gamma$ tracking filter with a noisy jerk as the maneuver model: frequency domain solution," IEEE Transactions on Aerospace and Electronic Systems, 1996, vol. 32, issue 3, pp. 1170-1174.
- [8] Kalata P. R., "The tracking index: A generalized parameter for $\alpha - \beta$ and $\alpha - \beta - \gamma$ target trackers," IEEE Transactions on Aerospace and Electronic Systems, 1984, vol. AES 20, issue 2, pp. 174-182.
- [9] Hoffman S. A. and Blair W. D., "The $\alpha - \beta - \gamma$ tracking filter with a noisy jerk as the maneuver model," IEEE Transactions on Aerospace and Electronic Systems, 1994, vol. 30, issue 3, pp. 925-927.
- [10] Sudano J. J., "The $\alpha - \beta - \gamma$ tracking filter with a noisy jerk as the maneuver model," IEEE Transactions on Aerospace and Electronic Systems, 1993, vol. 29, issue 2, pp. 578-580.

A Behavioral Approach to Estimate Area of Expertise for Cooperative Q-Learning: A Self Referenced Strategy

Ahmad Imanipour^{1,2}, Majid Nili Ahmadabadi^{1,2}, Babak N. Araabi^{1,2}

Masoud Asadpour¹, Roland Siegwart³, Aram Zaerpour¹

¹ Control and Intelligent Processing Center of Excellence, Robotics and AI Lab, School of ECE, University of Tehran, Iran

² School of Cognitive Sciences, Institute for Studies in Fundamental Sciences (IPM), Tehran, Iran

³ Autonomous System Laboratory (ASL), ETHZ, Switzerland

imanipour@ipm.ir, mnili@ut.ac.ir, araabi@ut.ac.ir, asadpour@ut.ac.ir, rsiegwart@ethz.ch

Abstract : Using each other's knowledge and expertise in learning – what we call cooperation in learning - is one of the major known methods to reduce the number of learning trials, which is crucial in real world learning applications. In situated systems, robots, even homogeneous ones, become expert in different areas due to being exposed to different situations and tasks. As a result, Area of Expertise (AOE) of any other agent must be detected before using its knowledge, especially when the exchanged knowledge is not abstract, or when a simple information-exchange is subject to misinterpretation, which is the case in Cooperative Q-Learning.

In this paper, we tackle the problem of AOE estimation from a behavioral viewpoint. At first, a Learning Agent (LA) develops a conception of self expertise by evaluating its own behavior at exposed situations. To this end, three different heuristic algorithms for a Q-learning agent are proposed. In each of these algorithms, the LA creates a table called Behavioral Self Evaluation Table (BSET) to record a quantity that may be perceived as a sign of expertness. The BSET is developed during the individual learning process. The LA utilizes the BSET to estimate its own AOE, while other agents do not necessarily have such a perception of themselves. Then, a classifier is designed and trained for the LA to discriminate AOE's of other agents. Proposed classifier uses Q-table based features as input, while utilizing BSET based AOE and non-AOE class labels for training. At this stage, the LA estimates others' AOE's, merely by looking at their Q-table based features. Extracted AOE's are merged in the last stage of the introduced cooperative learning method.

The proposed method is tested both in extensive simulations and in real world experiments using homogeneous mobile robots. The results show effectiveness of the introduced approach both in accurate extraction of AOE and in increasing the quality of the combined knowledge even when there is uncertainty and perceptual aliasing in applications and robots.

Keywords: Cooperative Learning, Area of Expertise, Knowledge Evaluation, Behavioral Self Evaluation Table, Q-Learning, Multi-Robot Learning.

چکیده: همکاری در یادگیری می‌تواند روند یادگیری در تیمی از عامل‌های یادگیرنده را سرعت بخشد. اما حتی برای عامل‌های همگن اجرایی کردن همکاری مشکلات فراوانی دارد و یکی از این مشکلات برآورد ناحیه خبرگی هر عامل است. در این مقاله به تخمین ناحیه خبرگی از یک منظر رفتارگرا پرداخته‌ایم. در این روش ابتدا هر عامل با تحلیل موقعیت‌هایی که در آن قرار گرفته به درکی از میزان و ناحیه خبرگی خود می‌رسد. به این منظور جدولی به نام جدول خود-ارزیابی رفتاری تعریف و حین یادگیری فردی، همگام با جدول Q، ایجاد می‌شود. سپس با استفاده از دانش عامل یادگیرنده از خبرگی خود نوعی طبقه‌بندی کننده طراحی و آموزش داده می‌شود تا به جدا کردن حالت‌های خیره و غیر خیره هر عامل دیگر بپردازد. به این ترتیب عامل یادگیرنده به تخمینی از ناحیه خبرگی سایر عامل‌ها در فضای حالت می‌رسد. روش پیشنهادی در شبیه‌سازی‌های فراوان و نیز به کمک آزمایش با ربات‌های متحرک ارزیابی شده و کارایی آن نشان داده شده است.

کلمات کلیدی: همکاری در یادگیری، ناحیه خبرگی، ارزیابی دانش، جدول خود-ارزیابی رفتاری، یادگیری Q، یادگیری چند رباته.

1- Introduction

Increasing the speed of learning and improving the quality of the learned knowledge are among the major goals of researches in machine learning. Along with devising faster learning techniques, creation of different methods of cooperative learning¹ has been the target of some researchers in the recent years. The main idea in cooperative learning methods is getting benefit from the existing knowledge in different agents through explicit and implicit exchange of learned rules, gathered information, etc. Consultation, voting, competition, guidance, imitation and information exchange are among the methods which facilitate and speed up the process of different learning methods.

Some of these methods are combined with reinforcement learning as it is a flexible, general, and semi-supervised learning method. It helps the LA to maximize its return by maintaining a balance between exploration and exploitation of the knowledge that the agent acquires from different sources and with different methods, ex. 000. Some of the mentioned methods, like imitation, advice taking, and consultation, require a mentor or a higher level of cognition or knowledge. In contrast, information-exchange-based methods are much less complicated and can be used by simple but similar agents. In this work we focus on information-exchange-based cooperative learning methods that are based on exchange of experiences or Q-tables among LAs 00. Studied agents are homogeneous in the sense of state-action representation.

The Q-learning 0 in its basic form is categorized as a slow method for large problems as LA must sufficiently search its state-action space. Some compact state-action representation methods, like fuzzy representation and hierarchical learning approaches (ex. 00) are proposed to reduce the size of Q-table. However, the convergence condition of the proposed learning methods is not clear yet. In addition, regardless of the state-action representation methods, agents might have different levels and Areas of Expertise due to being exposed to different conditions and tasks. Therefore, there is always a room for exchange of knowledge and information among the LAs. For example, in multi-agent systems, when reinforcement learning methods are exploited, the cooperative learning becomes

important, since the use of others' knowledge and experience lessens agent's need for individual experiments and even enhances its knowledge about inexperienced situations. Utilizing other agents' knowledge is more reasonable and reliable when agents are more expert.

To use other agents' knowledge and information, it is important to find their AOE as well as level of their knowledge. In some existing methods, like imitative learning, it is assumed that the mentor is fully knowledgeable; however, questions like *who* and *what* to imitate are two major unsolved problems in that field 0. Most of the time, this assumption is not valid in Q-learning agents as the agents might have non-equal chances for exploration of different areas of their state spaces.

Identification of AOE is not only important for cooperative learning but also for a class of learning methods, where the agents share some parts of their Q-tables in order to attain team-level cooperation. In such methods, ex. 0, the team may fail if the agents use each others' Q-values regardless of their domain and level of expertise.

In this paper we focus on identification of other agents' AOE for cooperative learning. To our best of knowledge, there is no other work on extraction of AOE at least for RL agents. In this research, based on a heuristic idea, first some algorithms are proposed for an agent to behaviorally measure its own expertness. In each of these methods, the agent creates a table to record a quantity that may be perceived as a sign of expertness. These tables are modified during the learning process. The LA is assumed to be aware of its own expertness through the aforementioned table, while other agents may not have such a perception of themselves. Based on concepts of pattern recognition, an algorithm is suggested so that the LA can evaluate others' AOE, using its own expertness and other agents' Q-table. Thus the LA is able to use the knowledge of more expert agents in an efficient manner. The detailed discussions are made for Q-learning agents. However, the main approach is general and can be extended to different learning methods. Fig. 1 shows a schematic diagram of the proposed method. Details are discussed in section III.

This research shows that proper understanding of expertness, and how to measure it, play a vital role in cooperative Q-learning. Indeed, the performance of cooperative Q-learning considering AOE may improve a lot compared to that of other existing

¹ Cooperation, cooperation in learning, and cooperative learning are used interchangeably throughout the paper.

cooperative Q-learning methods, where the AOE is not considered.

The rest of this paper is organized as follows: Section II discusses main cooperative Q-learning methods and clarifies the importance of extraction of AOE. The proposed method is introduced in Section III and the results of simulations and experiments are reported in Section IV. Conclusions and future works are given in the last section.

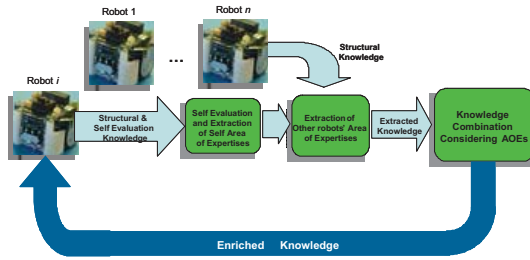


Fig. 1. A schematic view of the proposed method

2- Cooperative Q-learning

Cooperation in Q-learning can be performed in different manners. In *sensors sharing*, agents share their sensory data and each agent plays the role of a scout for others. In *episode sharing*, agents share (*state, action, reward*) triples. Since agent's experiences and knowledge are reflected in its Q-table, sharing Q-tables is a reasonable idea provided that the validity and the level of the encoded knowledge in Q-tables are considered. In *Simple Strategy Sharing* (SSS) or *Simple Averaging*, each agent learns in cooperation with its teammates through averaging their Q-tables:

$$Q_i^{new} \leftarrow \frac{1}{n} \sum_{j=1}^n Q_j^{old} \quad (1)$$

where n is the number of agents. SSS is not an appropriate method for cooperative learning when the agents' levels of expertise are different. Disregarding this fact, as in (1), may cause a misinterpretation that cooperative Q-learning is not beneficial for LAs. To solve the problem, *Weighted Strategy Sharing* (WSS) had been suggested in which each agent assigns weights to others by measuring their expertness and updates its Q-table as (2):

$$Q_i^{new} \leftarrow (1-\alpha_i) * Q_i^{old} + \alpha_i \sum_{j \in Expert(i)} \left(\frac{e_j - e_i}{\sum_{k \in Expert(i)} (e_k - e_i)} * Q_j^{old} \right) \quad (2)$$

where α_i is a weight that agent i assigns to its own Q-table, e_i is expertness of agent i and $Expert(i)$ is the set of agents that are more expert than agent i . Expertness is evaluated as a function of the history of the received rewards and punishments. Examples of utilized expertness measures may be found in Section III.A.3.

In (1) it is shown that WSS has a negative effect on speed and quality of learning when the agents have different AOE and their expertness values are measured using their entire rewards and punishments. Therefore, WSS is extended to consider agent's expertness at each state through evaluating the reinforcement signals it has received in that state. In addition, to specify the AOE, a method based on state transition is proposed. In this method, examining the entire Q-table, the agent detects all the states in which greedy action selection leads to achieving the goal. This technique is used as a gold standard and cannot be used for online cooperative learning.

Using WSS with state-level expertness measures is quite effective for cooperative learning among agents with different AOE's. However, calculation of state-level expertness requires recording the received reinforcement signals for each state, which is not practical. In addition, this method restricts the agents to cooperate in learning with those who have already saved their reinforcement history. To resolve this problem, in this paper we introduce a structural method for detection of AOE. By structural method we mean a method that uses a function of Q-values (like Entropy or MaxQ) and does not require observation of other agents' behavior or an explicit history of their learning.

3- The Proposed Method

In our proposed method, see Fig. 1, the agent that is willing to use other's knowledge, first evaluates itself through extraction of its own AOE by using a behavioral self-evaluation measure. Then, based on its self-evaluation, it constructs a Q-value-based meter to assess other agents' expertness and to estimate their AOE's. Finally, it combines experts' knowledge with its own. The main emphasis of this paper is on extraction of AOE.

Fig. 2 shows two main phases of AOE extraction; self evaluation and evaluating other agents.

In the first phase, the agent willing to evaluate other agents -Agent i in this figure- builds up a BSET while acting on the environment. Each cell of this table corresponds to one state of the agent and

shows agent's level of expertness at that state from a behavioral point of view. Then, using its BSET, the agent partitions its states into expert and non-expert classes. Members of these two classes are utilized as references for evaluating other agents in the second phase.

In the second phase, a classifier is used to evaluate other agent's knowledge and to extract their AOE's based on some exemplars. Each exemplar is composed of a feature vector and a class label; expert or non-expert. Since the only available information about other agents is their Q-tables in general, the employed features are Q-value-based. While, the references produced in the self-evaluation phase are *Agent i*'s states tagged as expert or non-expert. Therefore, an exemplar generator is used to establish the corresponding Q-value-based features of these states. The exemplar generator receives feature vectors that are derived from *Agent i*'s Q-table, along with the label of each state, which comes from BSET and produces a set of labeled feature vectors as exemplars. This data is used to train the *Agent i*'s classifier. The classifier gets *Agent j*'s states along with their features as input and classifies each state as expert or non-expert. The expert states form *Agent j*'s AOE from *Agent i*'s point of view. Each part of the proposed method is described in more depth in the sequel.

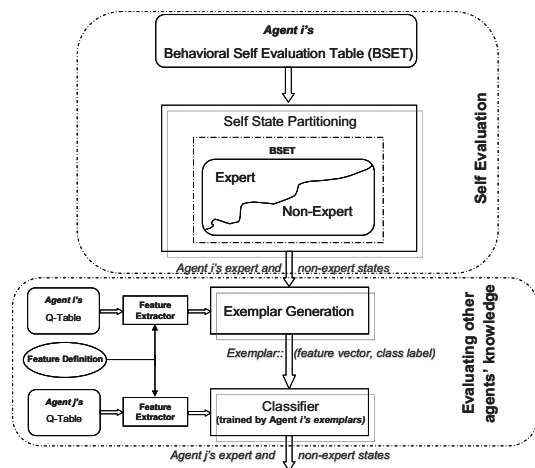


Fig. 2 The proposed method: self evaluation and evaluating other agents

A- Self-expertness evaluation

There is a consensus on the conceptual meaning of expertness; however, there are no general mathematical definition and evaluation measures for it. Based on findings in 0, here we argue that,

expertness is a relative concept in a society of agents and its emergence varies from one task to another. It means each agent should develop a mathematical conception as well as an evaluation measure for other agents' expertness based on its own task and knowledge representation method.

As our agents are reinforcement learner ones, their goal is maximization of received expected reward. They attain this goal through repetitive visit of states and balanced exploration and exploitation of decisions in addition to establishing a compromise between the instant and delayed rewards. Therefore, relative expertness of RL agents is embedded in their relative number of trials, dependable reinforcement signals, and the number of their sustainable success in achieving an explicit goal state –if such a goal exists. Therefore, to evaluate each other's knowledge, our agents calculate their relative expertness numerically based on one or a combination of this information.

In this paper, the definition of expertness is intended to measure and quantify the quality of agent's knowledge, not its behavioral characteristics, which require sufficient behavior observation as well as assumption of rationality in action selection. This way, the intended criteria can be used to evaluate other agents directly from their Q-tables. Therefore, among various definitions of expertness, the one that meets the following criteria at the highest level is opted:

- Its computation and modification is possible during the learning.
- It reflects the quality and domain of knowledge, which helps AOE extraction, and accommodates knowledge exchange.
- It has reasonable and realizable behavioral meaning.

As mentioned, maximizing the expected reward is the main goal of any RL agent. In some tasks the expected reward is maximized if agent reaches some specific states called goal states. Assuming explicit goal states simplifies the development of self-expertness evaluation methods; however, those methods are not applicable in general cases. Therefore, in the following three subsections we first introduce a general technique based on experienced states for self-expertness evaluation. Then we develop a counterpart method for the case that agent should reach some explicit goal states. Finally, a method which is based on agent's reward history is introduced.

The proposed methods use different information for expertness evaluation. These information are correlated to some extent, however. They represent different concepts. Therefore, the introduced methods refer to three different but partly related interpretations of expertness. As a result, the AOE that they identify can have non-overlapping areas. Availability of the required information and the intended application are two main criteria for selection of the most appropriate expertness evaluation method. However, fusion of these three methods, which is beyond the scope of this paper, can result in a more robust and comprehensive expertness evaluation.

A-1- Experienced States: Visit-table method

In this method, being experienced is taken as being expert. In other words, the more an agent explores a state, the more expert it becomes in that state. This interpretation of expertness is based on the assumption that, the agent in average makes its estimation of state's value more accurate, which means gaining more knowledge, at every visit. This assumption is true in RL methods, where the learning rate decreases as the learning proceeds and the share of the experience is increased in the agent's knowledge. Visiting each state, the agent selects one of its possible actions which explores or exploits a higher expected long term reward. The result, which is increase in the degree of awareness about the environment, can be exploited by other agents even if the agent does not behave rationally. From this viewpoint, the area of an agent's expertness and depth of its awareness can be formed in a *Visit-table*, $Visit(.)$, based on the number of visits of each state.

If all the visits of each state are counted regardless of the results, the table will be called absolute:

$$\text{Absolute) } s \xrightarrow{a} s' : \quad Visit(s) = Visit(s) + 1 \quad (3)$$

Highly experienced areas of Visit-table indicate the qualified knowledge of corresponding states of Q-table.

Although each visit affects Q-values, by defining some limitations on Visit-table computation, different special concepts can be inferred. For example, if the intended visits are those leading to increase the value of the selected action, a positive experience has taken place. Therefore:

$$\begin{aligned} \text{Positive) } s \xrightarrow{a} s' : \\ \text{if } (\Delta Q(s, a) > 0) \text{ then } Visit(s) = Visit(s) + 1 \quad (4) \\ \text{where } \Delta Q(s, a) = Q_{t+1}(s, a) - Q_t(s, a) \end{aligned}$$

Similarly, if Eq. (3) is limited to the actions which result in decreasing $Q(s, a)$, a negative experience has happened. Therefore:

$$\begin{aligned} \text{Negative) } s \xrightarrow{a} s' : \\ \text{if } (\Delta Q(s, a) < 0) \text{ then } Visit(s) = Visit(s) + 1 \quad (5) \\ \text{where } \Delta Q(s, a) = Q_{t+1}(s, a) - Q_t(s, a) \quad) \end{aligned}$$

According to Eq. (5), states with higher Visit-values are those in which improper actions are explored more. That means the agent has a good knowledge about the state-actions that result in negative experiences. This measure is useful for identification of states leading to lower return.

A-2- Learned States (A Path to Goal): L-table method

In goal-based Q-learning systems, achieving the goal is of high prominence and desirability. The consecutive selection of state-action couples which takes the agent to the goal shows the expertness of the agent from a behavioral viewpoint and brings remarkable successive rewards to the agent. If the agent has an understanding of the learned states, it is able to decide how to act and how to exchange knowledge with its teammates, that is to decide whether it requires some knowledge from the other agents. In this definition, learned state means the state from which there is a path to the goal in such a way that the path can be traversed via a specified policy. Therefore, the transition from one state to the state which has been learned before, can add this state to the sets of learned states. In other words:

$$\begin{aligned} \text{Initial value: } L(\text{GoalState}) = 1 \quad L(s \neq \text{GoalState}) = 0 \\ s \xrightarrow{a} s' \quad L(s) = L(s') \quad (6) \end{aligned}$$

To memorize the learned states, the agent constructs an *L-table*, $L(.)$, in which $L(s)$ is updated in parallel with updating the Q-table. As Eq. (6) shows, the entries of L-table for goal states are initialized to 1. In the environments in which the goal is already defined, the states associated with the goals are specified and the L-table is initiated at the beginning of learning and can be updated after the beginning. But in the problems in which there is no obvious goal or the goal is not known from the beginning, initialization should be postponed. To do so, after some trials, more valuable state(s) through searching the entire problem state space are determined as goal and then the L-table is initiated. This suggestion originates from the fact that the basis of reinforcement learning is to gather the maximum positive reinforcement signals.

In Q-learning, action selection is based on EEP (Exploration Exploitation Policy) approaches in which the emphasis is on the action that has had the highest benefit. Therefore, Eq. (6) is limited to the action with the highest value:

$$\text{If } \{s \xrightarrow{a^*} s' \text{ and } Q(s, a^*) = \max_{a \in \text{Actions Set}} Q(s, a)\} \\ \text{then } L(s) = L(s') \quad (7)$$

This way, L-table shows the learned states and their scope, if the agent exploits its knowledge.

This method is applicable when agent's goal is reaching an explicit state. In such case, agent is considered expert at a state if it has a sustainable experience in reaching the goal from that state. Similar to Visit-table method, L-table approach is independent from the value of reinforcement signal.

Random initialization of Q-table and high uncertainty in state transitions can result in premature change of L-table's cell. To solve this problem, the number of successful transitions to the goal from a state can be counted. Then the corresponding cell of L-table is changed to one when that counted number exceeds a certain threshold.

A-3- Gained Signals: E-table method

In *L-table* method, it is supposed that the goal exists and the agent rationally looks for it, while this is something beyond the basis of reinforcement learning in general form. In reinforcement learning, the movement of agent and reception of reinforcement signals forms knowledge, which is stored in and presented by Q-table. The decision-making and Q-table modification mechanism are based on the idea that at each state the agent estimates the expected value for each action during the learning progress. This estimation helps agent behave in such a way that it receives more reward and less punishment. From this viewpoint, "expertness of agents" can be measured by some kind of summation of received signals at each state 0. To do so, the sum of received signals for each state is stored in a table called *E-table*, $E(.)$, during the learning process, in order to be used in specifying the AOE during the cooperative learning process.

In this method, becoming expert is directly related to the history of received reinforcement signal. While each cell of E-table accumulates the gained reinforcement signal, higher values of its cell can be the result of getting a limited number of large reinforcements or receiving many small rewards.

Therefore, value of reinforcement signal should be considered in classification of states into expert and non-expert. However, as this method evaluates relative expertness, it is always true that agent is more (less) expert in states with higher (lower) *E-values*.

In computing *E-table* values, the type of signals (reward and punishment) can be considered in addition to their values. Their algebraic or absolute sum can be calculated separately or entirely. In the following equations, $E(s)$ is a criterion for estimation of agent's expertness at state "s" and $r_t(s)$ is the value of reinforcement signal received by the agent at the t^{th} visit of state s . Index t shows the length of summation interval that typically begins along with the learning process and continues to the last time when the agent has passed state s by selecting an action.

Normal (Nrm): According to this criterion, the agent's success and getting reward is the sign of expertness, while receiving punishment signals shows inexpertness, that is:

$$\text{Normal: } E(s) = \sum_{t=1}^{\text{last}} r_t(s) \quad (8)$$

Absolute (Abs): Either reward or punishment can enhance the knowledge of agent about the environment. Therefore, the level of expertness can be estimated by considering both of them:

$$\text{Absolute: } E(s) = \sum_{t=1}^{\text{last}} |r_t(s)| \quad (9)$$

Positive (Pos): According to this criterion, the experiences with no reward are ignored and just successful experiences are considered. The sum of positive reinforcement signals is computed as:

$$\text{Positive: } E(s) = \sum_{t=1}^{\text{last}} r_t^+(s) \\ r_t^+(s) = \begin{cases} 0 & \text{if } r_t(s) \leq 0 \\ r_t(s) & \text{elsewhere} \end{cases} \quad (10)$$

Negative (Neg): Just unsuccessful experiences are considered in this criterion. The more unsuccessful trials the agent has, the more expert it is in avoiding undesirable actions. This criterion is shown by the absolute sum of negative signals:

$$\text{Negative: } E(s) = \sum_{t=1}^{\text{last}} r_t^-(s) \\ r_t^-(s) = \begin{cases} 0 & \text{if } r_t(s) \geq 0 \\ r_t(s) & \text{elsewhere} \end{cases} \quad (11)$$

Above criteria have already been suggested to assign weights to the cooperative agents' knowledge on the entire Q-table or at each state in order to be used in cooperative learning by Strategy Sharing 0. The

direct use of Q-table values without any mapping in order to extract the AOE is not appropriate and efficient, since these values specify a better decision only when they are compared to each other.

In this research, the appropriateness of E-table method to perceive the expertise of each agent and evaluate others' expertise is studied and compared with that of L-table and Visit-table methods. Visit-, L-, and E-table are all examples of BSET in Fig. 2. Having one of the above mentioned tables, the agent can estimate the states in which it is expert. In case of L-table, class of expert states is evident by definition. In case of Visit- and E-table, different classification methods may be employed to categorize states into expert and non-expert classes based on their corresponding $Visit(.)$ or $E(.)$ values. In this paper we simply use median of $Visit(.)$ or $E(.)$ values as a threshold. Our experiments show that even this simple classification scheme works pretty well in self evaluation phase.

In the proposed methods, expertness is evaluated at the state level however; explicit regional expertness evaluation seems more robust. The reason behind state-level evaluation, instead of region-level, is twofold. First, in RL problems in general, there is no model available to detect neighbor states, and memorizing state transitions for this purpose is not practical. Second, expertness evaluation in Visit- and L-table methods is implicitly regional, as the neighbor states affect each other in Q-value updating. In other words, these tables are based on agent's behavior, which in turn is related to both the environment and agent's Q-table.

B- Estimating Expertness of Others

Knowing its AOE, now the evaluating agent should discriminate expert and non-expert states of other agents –or equally estimate their AOEs- for knowledge combination, see Fig. 1 and

Fig. 2. If all agents have created and updated their BSETs, no agent needs to estimate other's AOEs. In such a case, cooperative learning can proceed more accurately and efficiently, because agents can explicitly extract and exchange their AOEs. However, requiring all agents to have a BSET, especially in open systems –where new agents join and leave the system autonomously- is very restrictive. In addition, agents might use different BSETs and AOE extraction methods. In such cases, each agent needs to estimate other agents' AOEs from its own point of view.

In our approach, as an alternative, evaluating agent tries to estimate other agents' AOEs merely based on their Q-tables, along with its own BSET

and Q-table. In this paper, we tried to utilize a self evaluation metric to find a way to discriminate others' expert and non-expert states, just by looking at their Q-tables. To this end, we exploit the fact that if there are several known states as labeled samples (expert or non-expert), classification methods can be used to discriminate the AOE. This way, each agent knows its own AOE with the help of its BSET, and finds a Q-value-based reference, which helps finding others' AOEs.

In sum, estimation of other agents' AOEs is done in three main steps, see

Fig. 2. First, the evaluating agent forms a BSET and distinguishes its expert states from non-expert ones accordingly. Second, it calculates a set of features on its knowledge at each state –Q-cell in our case- and tags values of that features on its expert and non-expert states as such, respectively. In the last step, the agent computes the same features on other agents' Q-cells and employs a classifier to classify that feature values –and accordingly the related states- to expert and non-expert classes. This classifier utilizes the evaluating agent's tagged feature values as its training data.

We tested different functions for extraction of Q-value based features –i.e $entropy(.)$, $mean(.)$, $max(.)$, etc- and found $MaxQ(s)=max_a(Q(s,a))$ the most effective, while being less sensitive to the environment model and Q-table initialization. The reason that $MaxQ$ works better than the other candidates can be speculated in this way that, $MaxQ$ is agent's estimation of its highest expect reward in state s and it approaches its real value while agent gains more experience in the environment. Therefore, it is related to agent's expertness in state s .

More specifically, when the evaluating agent wants to use others' knowledge, it takes the following steps:

Reliably divides its own state space to expert and non-expert groups. This partitioning can be done utilizing one of the introduced BSETs (e.g. Visit-table) along with a suitable threshold like Median.

Computes some Q-table-based features for each state and labels them as expert and non-expert, based on the information obtained in the first step. Labeled features are used to form a classifier. In this paper, $MaxQ(s)$ opted as the feature of choice.

$$MaxQ(s) \leftarrow \max_a Q(s,a) \quad (12)$$

Specifies the core of the classification system using a conventional method such as K-Nearest Neighbors (KNN), Parzen or Neural Network classifier, and train the classifier using the labeled features obtained

in second step. That is, classifier is trained to discriminate expert and non-expert states as similar as what is obtained through BSET. In this paper, we opt to use a Parzen classifier.

Estimates AOE of other agents by computing the same Q-table-based features on each state of their Q-tables and feeding them to the classifier. The AOE of the other agent is extracted as the output of the classification process.

In the experiments, it is assumed that the reward function is the same all over the state space. This assumption could be relaxed to some extent as the usage of a classifier brings in robustness against minor change of reward function from one region of states to another. Therefore, the more relaxed assumption is that the difference of state regions' reward functions is limited. When regions have very different reward functions, self-expertness evaluation and estimation of other's expertness should be done region by region. The evaluating agent becomes eligible for evaluation when surpassing a minimum number of experiences. Having the same or sufficiently similar reward functions for all regions provides the chance of having experience in one region and evaluating others in different areas for the evaluating agent.

C- Knowledge combination

As explained, the existing methods for cooperative learning 0,0 not only lose their power in enhancing agents' knowledge, but also can have negative effect on each individual's knowledge when agents have different AOE's 0. Therefore, our agents extract others' AOE's from their perspective first. Then each agent combines other's knowledge with that of itself accordingly. Preference in knowledge selection for combination is given to knowledge at AOE's.

There are different methods for merging Q-tables 0; however, in this paper we use a simple method, called Expertness-based Strategy Sharing (ESS), in order to show effect of extraction of AOE more explicitly. Having extracted its own and other agents' AOE, each robot combines its own Q-table with that of the other ones to create a new Q-table (see Fig. 3).

As Fig.3 shows, when merging, each robot keeps its Q-values in the states it evaluates itself expert and copies either Q-cells of non-overlapping AOE's of others or the Q-cell of the most expert agent. Q-values of the remaining states are the average of agents' corresponding Q-values.

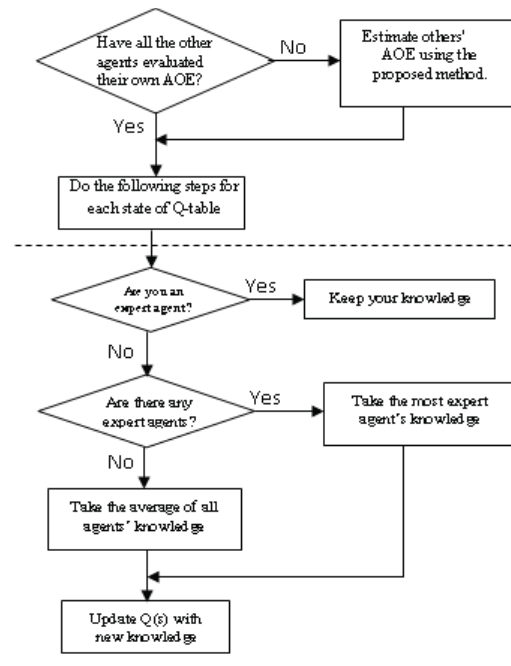


Fig. 3. The flowchart of knowledge combination algorithm based on Expertness-based Strategy Sharing (ESS).

4- Simulation and Experimental Results

We present both simulation and experimental results to show the benefits of cooperative learning, importance of extraction of AOE, and performance and applicability of our proposed method in learning tasks with and without explicit goal states in deterministic and stochastic environments.

The proposed approach is a general framework however, in order to make the results more visually presentable; in the simulations we use a maze-like environment with three distinct goals. In real robot experiments no goal state exists and the robot is expected to learn a continual behavior.

For the Parzen Classifier the window size is set to 0.3 in simulations and 8 in experiments. These numbers are derived through gradual increments of window size in search for the least misclassifications.

A- Simulation results

An environment with 31x21 squares is defined– see Fig. 4. There are three goal spots –one in the middle of the lower part, one in the center, and one in the middle of the upper part– and some obstacles. Such environment makes it possible to force the agents to gain different AOE as well as makes the presentation of the results easy especially via images. By forcing the agents to gain different AOE we mean

initializing agents with higher probabilities in the desired area. Doing so, the agents explore and learn more in the desired area.

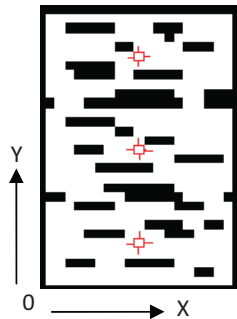


Fig. 4. A schematic view of the simulation environment. The goals are shown by crosses

At each state (s_t), the agent selects one of four actions (a_t), up, down, left or right, with a probability which is specified by Boltzmann function:

$$p(a_t | s_t) = \frac{e^{Q(s_t, a_t)/\tau}}{\sum_{k \in \text{Actions}} e^{Q(s_t, a_k)/\tau}} \quad (13)$$

Where $\tau=1.5$ is the temperature. The agent receives +1 reward by achieving the goal, -1 punishment by hitting the obstacles and -0.1 in other cases. The Q-values are updated according to TD (0) Q-learning:

$$Q(s_t, a_t) \leftarrow (1-\alpha)Q(s_t, a_t) + \alpha(r + \gamma V(s_{t+1})) \quad (14)$$

Where the learning rate α is 0.8, the discount parameter γ is 0.9 and $V(s) = \max_{a \in \text{Actions}} Q(s, a)$.

Simulations are done both in deterministic and non-deterministic environments and Q-tables are initialized with zero and random values between +1 and -1. In order to evaluate our approach, we have performed many simulations under different conditions -like different levels of uncertainty in the environment and Q-table initialization- and with agents with differences in areas and levels of expertise. The results show the effectiveness of our approach in all cases however; due to space limitation we report details of just a set of representative results. General outcomes are also reported.

A-1- Extraction of AOE

A-1-1- Deterministic environment

If agents choose the starting points in specific areas with a higher probability, it is expected that they

have different AOE with the equal number of trials. That is, each agent will gain more knowledge about the states located at its own area of activity. Fig. 5 to Fig. 8 show the extracted areas in which each agent has gained knowledge by itself through Eqs. (4), (5), (7), and (9).

Note that the cells of L-table are set to either 0 or 1 and represented with 2-level (black and white) pictures, but the cells of other tables (Visit-table or E-table) may get different values. Therefore, their representation needs a proper mapping to multi-level pictures, or they can be shown with 2-level pictures by choosing a suitable threshold which can be the mean or median of the values.

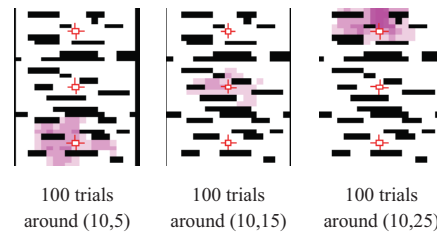


Fig. 5. AOE based on positive experiences (Visit-table) of each agent using Eq. (4)

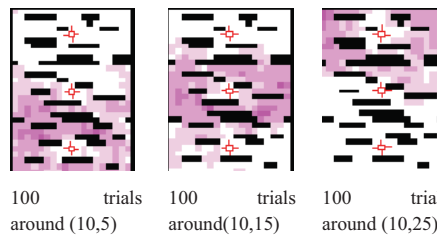


Fig. 6. AOE based on negative experiences (Visit-table) of each agent using Eq. (5)

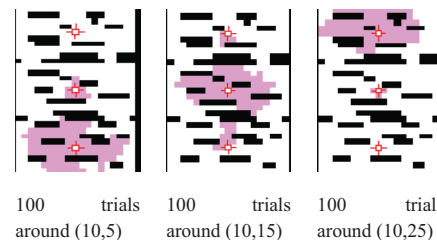
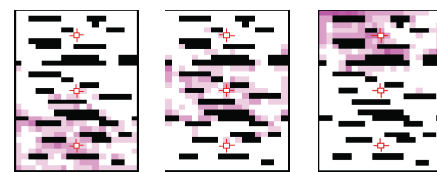


Fig. 7. AOE based on learned states (L-table) of each agent using Eq. (7)



100 trials around (10,5) 100 trials around (10,15) 100 trials around (10,25)

Fig. 8. AOE based on Abs received signals (E-table) of each agent using Eq. (9)

Nrm -Eq. (8)- leads to similar results to Fig. 8. By using the Neg criterion -Eq. (11)-, the states around the goal become empty and by using Pos criterion -Eq. (10)- only the states around the goal are marked. Note that changing the value of reinforcement signals has observable effects on the strength and scope of the results for Eqs. (8) to (11), while the output of L-table and Visit-table are more value independent.

If agents explore the environment similarly but with different number of trials, it is expected that each agent's AOE is different from others'. Fig. 9 shows the results using Eqs. (4) and (7) when the agent has started learning with random initial Q-values between +1 and -1.

As expected, the more trials an agent has, the larger its AOE becomes. It means that, at a higher number of trials, the set of states in which the agent can reach the goal by greedy action selection (right side) and also its number of positive experiences (left side) are increased.

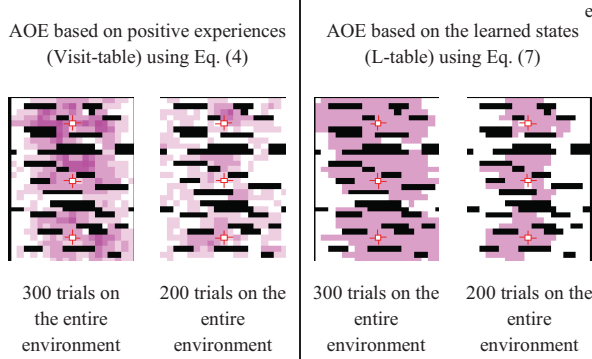


Fig. 9. Difference in AOE because of different number of trials when Q-tables are initialized randomly.

Fig. 10 shows how an agent having its own Visit-table -positive experiences in this case- has been able to properly estimate two other agents' AOE when all have the same number or trails. Similar results are obtained using L-table and E-table with some sparse states when using E-table.

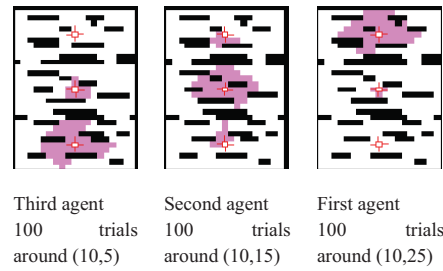


Fig. 10. The first agent extracts its own AOE and that of others' based on positive experiences (Visit-table) -Eq. (4)- and MaxQ

Now we test our methods when the agents have different numbers of trials. In Fig. 11, the agent uses its Visit-table and MaxQ to estimate others' AOE that have been trained in different numbers of trials. Using L-table leads to similar results.

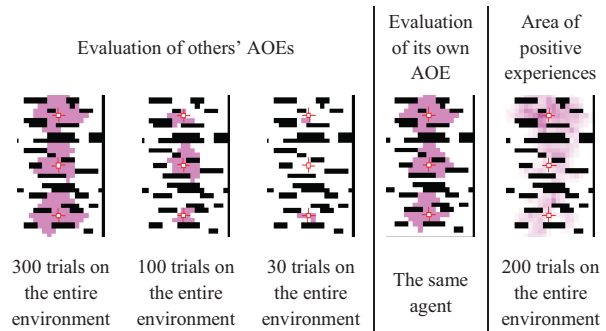


Fig. 11. The right side agent extracts its own AOE and that of others' based on positive experiences (Visit-table) -Eq. (4)- and MaxQ

To evaluate applicability of the suggested methods, a comparison is made between the AOE an agent perceives of its own using L-table or Visit-table with the area estimated by another agent. In

Table 1, the number of states that both agents evaluate the same and different are shown. In this table, the evaluating agent runs 200 trials on the entire environment and the self evaluating one runs different number of trials.

Table 1. Evaluation of AOE by the 200-trial agent in comparison with each agent's perception of its own expertise

	L-table		Positive Visit-table		Negative Visit-table	
	Equal	Different	Equal	Different	Equal	Different
The 30-trial agent	426	4	426	4	169	261
The 100-trial agent	408	22	392	38	234	196
The 200-trial agent	356	74	363	67	318	112

The 300-trial agent	274	156	358	72	351	79
---------------------	-----	-----	-----	----	-----	----

A-1-2- Stochastic environment

One important challenge in cooperative reinforcement learning is performance of the proposed solutions in nondeterministic environments. In this research a probabilistic state transition model is used to simulate such environments, see Table 2. This distribution is the same for all states.

Table 2. State transition probability

Action \ Result	Up	Down	Left	Right
Up	0.90	0	0.05	0.05
Down	0	0.90	0.05	0.05
Left	0.05	0.05	0.90	0
Right	0.05	0.05	0	0.90

Fig. 12 shows the areas that each agent has individually extracted based on positive Visit-table -Eq. (4)- and L-table -Eq. (7)-. It can be seen that not only the area of activity and detected AOE are similar, but also the results of using L-table and Visit-table are compatible.

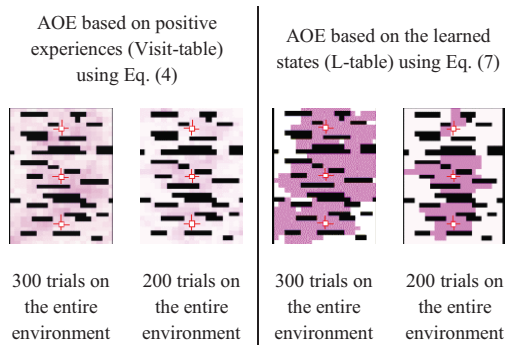


Fig. 12. Self extraction of AOE in a nondeterministic environment. Agents explore the whole environment.

Fig. 13 shows that the attained results of different experiments are still appropriate estimations of reality. Images of evaluating others' expertness are comparable to the same agents' perception of self expertness which is shown in Fig. 12.

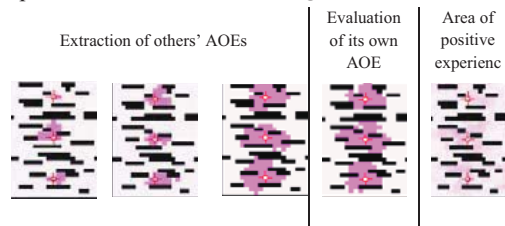


Fig. 13. The right side agent extracts its own AOE and that of others' based on positive experiences and MaxQ in a nondeterministic environment.

A-2- Knowledge combination

At the beginning of each group of simulations, four agents learn independently. Then, all Q-tables are combined based on ESS method using expertness, Visit-tables or L-tables of the agents. The results are compared to SSS and WSS with Abs expertness criterion -Eq. (9)- in terms of 'State Transition test' 0 and Euclidian distance between obtained Q-tables and that of a fully learned agent. State Transition test is a behavioral gold standard in which the states in those greedy actions leads to the goal state are counted. Euclidian distance of Q-tables is used as a structural measure to show the difference in agents' knowledge that is coded in Q-tables¹.

A-2-1- Deterministic Environment – Zero Initial Value

In the first group of experiments, the initial value of Q-table has been set to zero and the agents have been forced to choose the start point in a specific area.

Fig. 14 shows each agent's AOE under the State Transition test. The combined knowledge of these 4 agents by the agent that runs 100 trails around the upper goal is presented in Fig. 15.

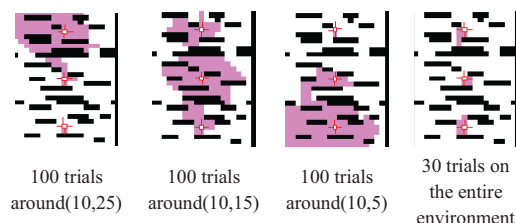


Fig. 14. Capability of agents running different number of trials to reach the goals in a deterministic environment (zero initial value)

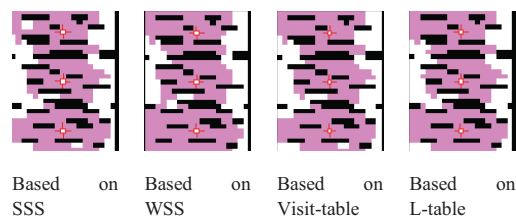


Fig. 15. Comparing the capability of the combined knowledge to reach the goals in a deterministic environment (zero initial value)

¹ We call this distance difference in stored knowledge.

Since the Q-tables are initialized by zero, combining them using SSS and WSS does not have a destructive effect on agents' behavior outside AOE. Therefore, as Fig. 15 indicates, all methods have similar results in terms of reaching the goals although the resulted knowledge has different magnitudes, see Fig. 16. It can be seen that firstly, the knowledge combination methods are able to use agents' knowledge and secondly, our proposed method is the most efficient one.

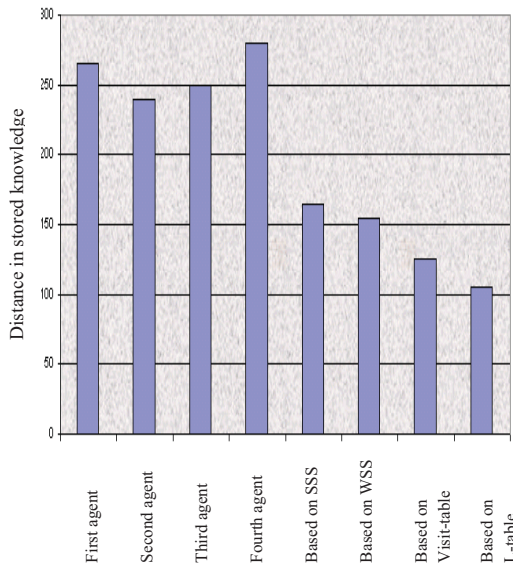


Fig. 16. The distance between the stored knowledge in each Q-table to the knowledge of a completely trained agent in a deterministic environment (zero initial value)

A-2-2- Deterministic Environment – Random Initial Value

Fig. 17 and Fig. 18 present the same results when the Q-tables are initialized with random values. According to Fig. 18, the knowledge combination methods based on Visit-table and L-table noticeably extend the AOE but, strategy sharing methods -WSS and SSS- have decreased the agent performance in reaching the goal. Hence, it may be thought that knowledge combination is not useful at all when using SSS or WSS methods. But, more accurate study and measuring the knowledge quality, see Fig. 19, show that the knowledge combination is generally useful during the learning from the knowledge quality point of view.

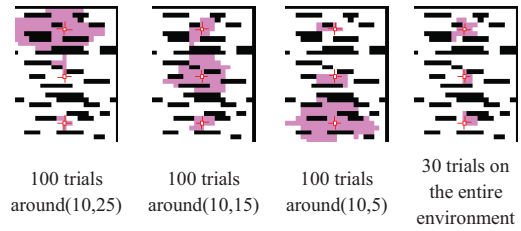


Fig. 17. The capability of each agent to reach the goal in the deterministic environment (random initial value)

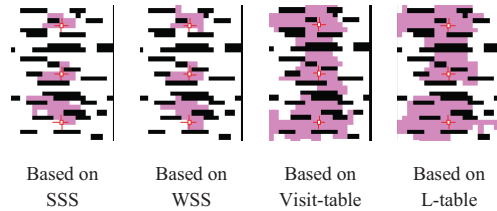


Fig. 18. Comparing the capability of combined knowledge to reach the goal using different methods in the deterministic environment with random Q-value initialization

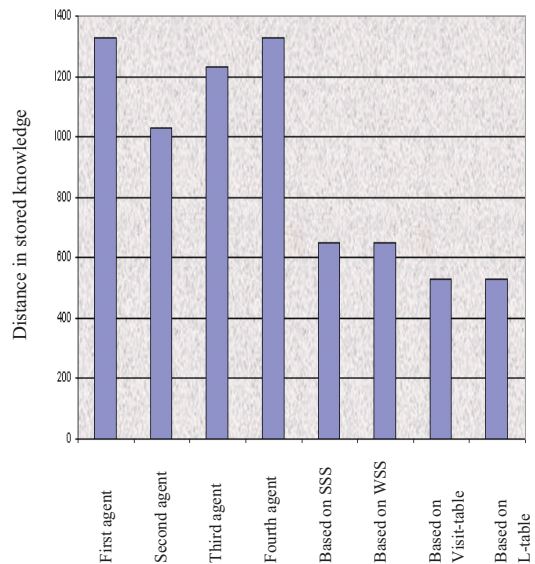


Fig. 19. The distance between the stored knowledge in each Q-table to the knowledge of a completely trained agent in the deterministic environment (random initial value)

In Fig. 19, it is clear that for all knowledge combination methods, especially for our proposed methods, the difference of combined Q-tables and the reference table is remarkably less than the others. In other words, if learning continues after knowledge combination, it is expected that the agents reach the final result faster than when continuing with individual learning.

A-2-3- Nondeterministic Environment – Zero Initial Value

Studying the same problems and methods in nondeterministic environment¹ shows that the preceding analyses and results are still more or less valid, although the positive effect of the proposed methods decreases when the environment become very nondeterministic. The distance of Q-tables and the one of a fully learned agent is shown in Fig. 20.

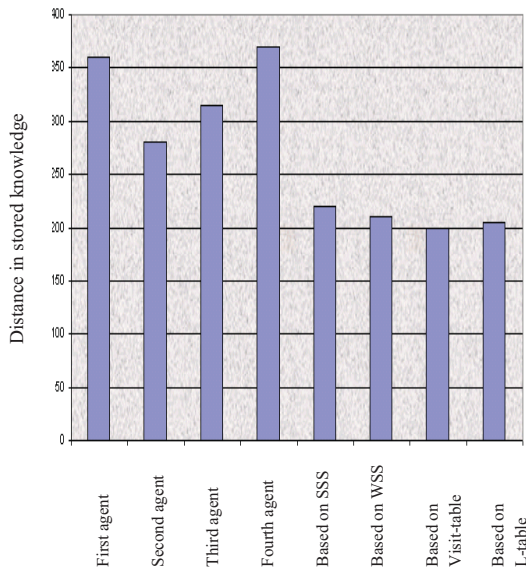


Fig. 20. The distance between the stored knowledge in each Q-table and a the one of fully trained agent in a nondeterministic environment (zero initial value)

A-2-4- Nondeterministic Environment – Random Initial Value

State transition matrix and Q-value initialization are as in Sections IV.A.2.2 and IV.A.2.3. Random initialization of Q-tables, as in Section IV.A.2.2, limits the success of WSS and SSS in State Transition test.

$$P(a_i | x) = \frac{Q(x, a_i) + 1}{|Actions| + \sum_{k \in Actions} Q(x, a_k)} \quad (16)$$

¹ The state transition probabilities in the rest of experiments are same as Table 2.

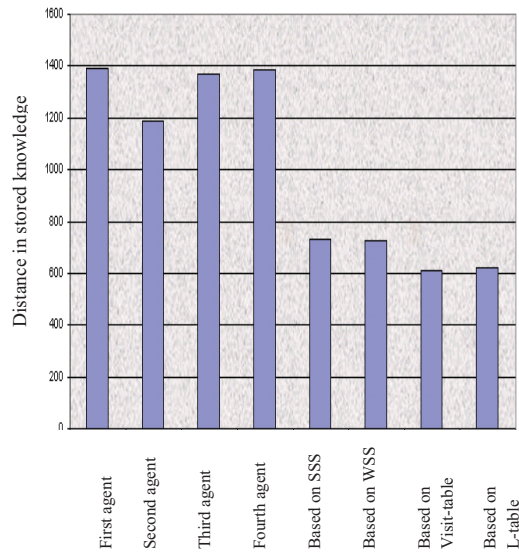


Fig. 21. The distance between the stored knowledge in each Q-table to the knowledge of a completely trained agent in a nondeterministic environment (random initial value)

The distance of Q-tables to the knowledge of the totally learned agent (Fig. 21) shows that the results gained by the knowledge combination methods are much better than individual knowledge of each agent. Our proposed method is again the best among them.

B- Experimental results

A set of experiments on two Alice micro-robots 0 is performed to test the applicability of the proposed method in real world applications. *Alice* is a small (22 x 21 x 20 mm) and light weight (5 gr.) autonomous mobile robot (see Fig. 22). It has two bi-directional watch motors for locomotion (up to 40 mm/s), four active infrared proximity sensors located on the Front (F), Back (B), Front Left (FL) and Front Right (FR) of its body, a PIC16F877™ micro-controller (@1 MHz) with 8K x 14-bit words flash program memory, 368 x 8-bit RAM data memory, and 256 x 8-bit EEPROM, and an IR TV remote receiver for communication. It has a NiMH rechargeable battery which provides autonomy of up to 10 hours.

A compact version of Q-learning has been developed for Alice 0. In this version, Q-values are unsigned integers and have a minimum value of zero. Actions are selected based on Roulette Selection, that is:

where $|\text{Actions}|$ is size of the action set. Policy is update as following:

$$Q(x, a) \leftarrow Q(x, a) + r + f(x, a, y)$$

$$f(x, a, y) = \begin{cases} \gamma(V(y)) & \text{if } Q(x, a) < \theta_Q \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

where r is the reinforcement signal, θ_Q is a threshold, $V(y)$ is the value of the best action in state y , and f is the “discount function” that depends on $V(y)$ and is implemented via conditional operators (e.g. if $V(y) > 32$ then $\gamma(V(y)) = 2$).

Alice observes 16 distinct states using its discretized infrared sensors. Each state shows the relative location of the observed obstacles to the robot. For example, Alice is facing toward the corner when its state is $(B=0, F=0, FL=1, FR=1)$. Due to its limited sensing capability it has an intrinsic perceptual aliasing which makes its learning a challenge. It has also three actions, Forward (FF), Turn Left (TL) and Turn Right (TR). It means, Alice has a 16×3 Q-table.

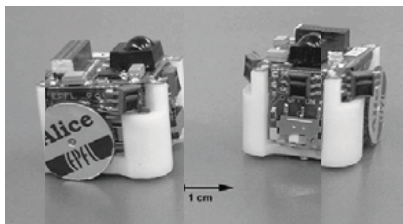


Fig. 22. Alice micro-robots.

We have two main choices to make two Alice robots expert in different domains. The first option is selection of two different and non-fully overlapping tasks. The second one is manipulating the robot’s environment or sensory system such that each robot experiences a subset of its states. We choose manipulation of sensory system because of having limited computational power and memory storage on Alice.

In the individual learning phase, we set $FL=0$ and $FR=0$ for the first and the second robot respectively. It means the first robot (*AliceBL*) is blind to the obstacles on its front left and the second one (*AliceBR*) cannot see the objects at its front right direction¹. As a consequence, the first and the second robots do not experience states $(B=*, F=*, FR=*, FL=1)$ and $(B=*, F=*, FR=1, FL=*)$ respectively. Therefore, each robot observes the world in 8 states

and the world is less observable and more uncertain for the robots. Also, the learned rules for observed states are affected by the robot’s experiences in the projected states. Fig.23 shows one set of the world states as sensed by *AliceBL*. A similar perceptual aliasing exists for *AliceBR*.

We used two different environments –namely a narrow hallway and a plus-shape maze- and performed separated sets of experiments in each; see

Fig. 24 and Fig. 25. It is expected that our method shows its strength in the narrow hall environment more as the engineered perceptual problem affects the robots’ in this environment more.

In individual learning phase, each robot is placed separately in the environment with Q-cells initialized at 5. The robot selects an action every 200 ms. The robot gets +1 reward when it goes forward and does not get closer than 1 cm to the obstacles. It gets -1 when it runs too close to the obstacles, and receives 0 in other cases. Having these punishments and reward, the robot is expected to move forward as much as possible and avoids obstacles with minimum number of turns.

After 17 minutes of individual learning by *AliceBL* and *AliceBR*, their Q- and Visit-tables are collected. Then, each robot extracts its own and the other one’s AOE and merges their Q-tables using the discussed method.

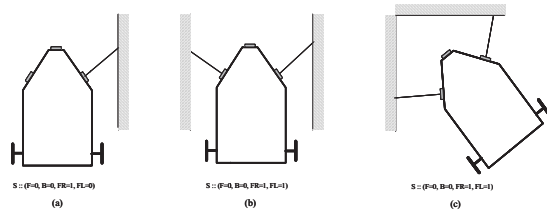


Fig.23. AliceBL observes all three states the same.



Fig. 24. Alice in a narrow hallway.

¹ In fact we used the same robot for the experiments for the sake of being fair in the following evaluations however; we use two different names for the robots for simplification of discussions.

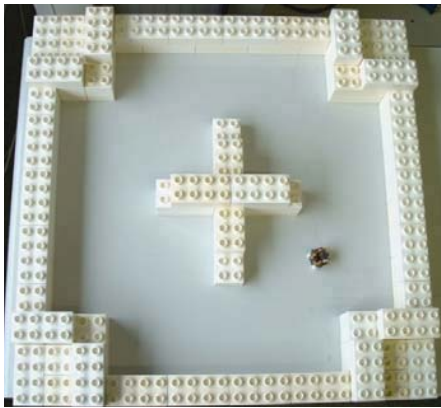


Fig. 25. Alice in a plus-shape maze.

In the test phase all sensors' status are returned back to normal. The robot with fully functional sensors is tested using *AliceBL*, *AliceBR*, Simply Averaged 0, and two Q-tables merged by *AliceBL* and *AliceBR*. The robot runs 20 minutes with each Q-table in the environment without learning and sum of the received reinforcement signals is stored every 15 sec in EEPROM. Then, the gained reinforcements are downloaded and analyzed.

Table 3 shows the average reinforcement signal that the robot has received during every 15 sec. using the five mentioned Q-tables.

As the results show, the proposed method works much better than single agents and simple averaging in the narrow hallway. The introduced approach also act much better than WSS 0 in the narrow hallway as the agents have different AOE while WSS cannot detect this fact and just copies or calculates a weighted averages of Q-tables.

SSS and our approach have similar results in plus-shape maze. The reason for this observation is that, the free area in plus-shape maze is much wider than free space in the narrow hallway. This means both robots have a higher chance of observing no obstacle. In addition, the resulted Q-tables show that *AliceBL*'s AOE is a subset of *AliceBR*'s. Moreover, because of practical considerations Q-tables are not initialized randomly. Therefore, SSS acts very similar with our approach when it is employed by the robot with less accurate knowledge, that is *AliceBL*.

Table 3. Average reward gained during test phase using different Q-tables.

Robot and method Environment	<i>AliceBL</i>	<i>AliceBR</i>	Simple Averaging (SSS)	Proposed method by <i>AliceBL</i>	Proposed method by <i>AliceBR</i>
Narrow hallway	-28.63	-0.93	-0.02	7.13	7.83
Plus-shape maze	1.71	18.84	38	37.58	41.553

Numbers in

Table 3 show that *AliceBL* does not perform as well as *AliceBR* does. The major reason for this observation is that, *AliceBL* and *AliceBR* sense $S1=(0,0,0,1)$ and $S2=(0,0,1,0)$ respectively as $S0=(0,0,0,0)$. As a consequence, their experiences at these states affect Q-values of $S0$. This effect is more for *AliceBL* robot in the reported result as *AliceBL* has 15% less chance to choose *FF* when there is no obstacle around the robot in the narrow hallway. As $S0$ is encountered more often and *FF* is the rewarding action in that situation, *AliceBL* gets less reward compared to *AliceBR*. In addition, note that the robots have a myopic sense of the environment and the ratio of the robot's width to that of the environment is about 0.3 and 0.1 for the narrow hallway and the plus maze respectively. Therefore, turning right or left at $S0$ increases the probability of getting too close to the walls and getting punishments in the next steps. The merged Q-table by *AliceBL* has the same problem at $S0$ as it considers itself expert in that state. But, the chance of getting punishment in the next states is decreased because *AliceBL* gets correct knowledge from *AliceBR* at states $(*,*,0,1)$ and gains a more dependable knowledge by simple averaging of Q-cells at the states that both robots are not expert.

5- Discussions

By taking a closer look at the introduced methods, analytical discussions, the experimental results, and the extensive simulation studies, the following points can be observed and concluded:

- General results show that all proposed methods can detect AOE when Q-tables do not contain false knowledge and the environment is not highly stochastic. Positive Visit-table and L-table retain their power in AOE detection even when some Q-values are not correct. The results also show that AOE detected by positive Visit-table matches the one extracted using L-table. Detected AOE by means of L-table is smaller than positive Visit-table for low number of trials because the agent has gained more positive experiences in some states but it might have not reached the goal from there yet.
- Obtained but not reported simulation results show that AOE detected using negative Visit-table is a complementary of the one extracted by means of positive Visit-table. This outcome complies with the mathematical definition of positive and negative expertness.
- L-table works well, even when the Q-tables are initialized randomly and Visit-table perform acceptable, but it might make some

mistakes in the initial stages of learning. The number of mistakes is reduced when the learning proceeds.

- The area detected by Visit-table or L-table method is also detected by E-table but E-table have some extra sparse states detected as expert. This situation is due to the fact that these measures act very locally based on instantaneous rewards while L-table and Visit-table implicitly consider neighbor states and discounted reward as well. The number of these sparse states depends on the value of reinforcement signal and selected thresholds.
- In the same number of trials, AOE extraction methods are more conservative in more stochastic environments. It is natural as an agent itself needs higher number of trials to learn a stochastic task and its expertness grows slowly. In other words, it should be noted that, increase of uncertainty in the environment results in slower learning and flatter distribution of Q-values. Therefore, estimation of other agents' AOE becomes less accurate especially at lower number of experiences. As a consequence, effectiveness of the proposed method decreases by increase in uncertainty in the environment; however, it is much better than other methods yet.
- The results show that an agent's self evaluation and its evaluation by other agents are well matched using positive Visit-table and L-table. The results show that these two evaluations are closer if the evaluating agent is an expert or is more expert than the evaluated agent. In other words, the more expert an agent is the fewer mistakes in extraction of other's AOE. Our method in estimation of others' AOE is based on their structural knowledge not their behavior. Therefore, as the knowledge in RL agents is value-based, the reward function affects our approach when regions of state-space have different reward functions and the evaluating agent is not aware of this fact. The usage of a classifier brings in robustness against limited change of reward function from one region to another— or similarly from one agent to another. However, tolerable bound on possible changes should be studied.
- Results of knowledge combination in the simulations and the experiments indicate that AOE-based knowledge combination

outperforms single agent learning, SSS and WSS in all cases in terms of behavioral and structural measures. Applied structural measure is the difference of Q-tables with that of a fully learned agent. The behavioural measure is the number of states in which the agent can reach the goal by its best action for goal-reaching tasks or the average of gained reward for the other missions. The results show that benefits of using AOE-based method are more distinct when some Q-values are wrong- for example due to Q-table initialization or perceptual aliasing.

In sum, it can be generally concluded that the proposed methods are successful in representing agent's expertness however; L-table and positive Visit-table show the AOE better under various conditions. For self evaluation we propose using L-table when there is a goal state in the environment and positive Visit-table when the goal state does not exist or its existence is not clear. In addition, we suggest starting self evaluation after some initial learning trials when using Visit-table for agents with random Q-table initialization.

Here we used a scalar feature $-\text{MaxQ}$ - and a Parzen classifier for extraction of other agents' AOE. Feature vectors can be used to include more information in the feature space in addition to using other appropriate classifiers.

6- Conclusion and Future Works

In this paper we introduced an approach to extract other agents' AOE for cooperative learning using just their Q-tables. In this method, learning robot uses a behavioral measure to form a BSET and evaluate itself accordingly. The outcome of this process is a set of states where the agent is expert in. This set is called AOE. BSET is used then along with a Q-table-based feature to extract other robots' AOE using a classifier. Extracted AOE are used for cooperative Q-learning.

The initial results showed that the existing cooperative Q-learning methods result in partially incorrect and inaccurate knowledge, when difference in AOE is ignored, and particularly when the Q-tables are initialized randomly. In contrast, simulation and experimental outcomes indicated that the proposed cooperative learning is very effective in partially solving two major challenges in real world applications of Q-learning. The challenges are reducing the number of learning trials and increasing

the quality of the learned knowledge, when different sources of knowledge are present and appropriate knowledge for adoption is identified through extraction of AOE.

The introduced method uses a behavioral self-evaluation measure along with a simple function of other agents' Q-table to extract their AOE. Making the process of reliable knowledge identification less dependent to self-evaluation measures is one of our future research plans. Extension of the presented notions to other learning methods and test of our approach in more complex robotic systems are among our next research steps.

The presented idea is general however; detailed definitions are developed for discrete state-action spaces. Our preliminary efforts in extension of the presented approach to fuzzy RL systems are promising and a more extensive study to adopt the presented AOE extraction methods to continuous RL approaches is underway.

References

- [1] ME. Lopez, R. Barea, Lm. Bergasa, MS. Escudero, "A human-robot cooperative learning system for easy installation of assistant robots in new working environments", *Journal of Intelligent & Robotic Systems* 40 (3): 233-265, 2004.
- [2] B. Price and C. Boutilier, "Accelerating Reinforcement Learning through Implicit Imitation" in *Journal of Artificial Intelligence Research*, Vol. 19 pp. 569-629, 2003.
- [3] L. Nunes, E. Oliveira, "Cooperative Learning Using Advice Exchange", in *Adaptive Agents and Multi-Agent Systems*, pp. 33 – 48, April 2003
- [4] M. Tan, "Multi agent Reinforcement Learning: Independent vs. Cooperative Agents", *Proc. 10th Int. Conf. on Machine Learning*, Amherst, MA, pp. 330-337, Jun. 1993.
- [5] M. Nili Ahmadabadi and M. Asadpur, "Expertness based Cooperative Q-learning", *IEEE Trans. System, Man, and Cybernetics Part B*, vol. 32, no. 1, pp. 66-77, 2002.
- [6] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, MIT Press, Cambridge, MA, 1998.
- [7] T.G. Dietterich, "Hierarchical reinforcement learning with the MAXQ value function decomposition", *Journal of Artificial Intelligence Research*, 13, 227-303, 2000.
- [8] L. Jouffe, "Fuzzy inference system learning by reinforcement methods", *IEEE Trans. on Systems, Man, and Cybernetics, Part C*, Vol. 28, No. 3, 1998.
- [9] K. Dautenhahn & C. L. Nehaniv (Eds.), *Imitation in Animals and Artifacts*, MIT Press, 2002.
- [10] Claude F. Touzet, "Distributed Lazy Q-learning for Cooperative Mobile Robots", *International Journal of Advanced Robotic Systems*, Volume 1 Number 1, pp. 5-13, 2004.
- [11] Z. Ren and A. B. Williams, "Lessons learned in single-agent and multiagent learning with robot foraging" in *IEEE Int. Conf. on Systems, Man and Cybernetics*, pp. 2757- 2762, 2003.
- [12] S. Mastour Eshgh and M. Nili Ahmadabadi, "An Extension of Weighted Strategy Sharing in Cooperative Q-learning for Specialized Agents", *Proc. 9th Int. Conf. on Neural Information (ICONIP'02)*, Singapore, pp.106-110, Nov. 2002.
- [13] B. N. Araabi, S. Mastour Eshgh, and M. Nili Ahmadabadi, "A Study on expertness of agents and its effects on cooperative Q-learning" Accepted for publication in *IEEE transaction on System, Man, and Cybernetics, Part B*, Mar. 2007.
- [14] G. Caprari, "Autonomous Micro-robots: Applications and Limitations", Ph.D. Thesis, EPFL, Switzerland, 2003.
- [15] M. Asadpour and R. Siegwart, "Compact Q-Learning Optimized for Micro-robots with Processing and Memory Constraints", *Journal of Robotics and Autonomous Systems*, vol.48, pp.49–61, 2004.
- [16] A. Imanipour, M. Nili Ahmadabadi, and B. N. Araabi, "Estimation of area of expertise for cooperation in fuzzy Q-learning" in *12th International CSI Computer Conference (CSICC'2007)*, Tehran, Iran, , Feb. 2007.

The application of non-cooperative Game theory in B2B electronic market

Sahar Sohrabi¹, Mohammad Jafar Tarokh²

¹Post graduate student, shr.sohrabi@gmail.com

²Associate Professor, mjtaroKh@kntu.ac.ir

Department of Industrial engineering, K.N.Toosi University of technology

Abstract: B2B electronic market facilitates the transactions among businesses. Unfortunately, most of these markets couldn't generate the expected profits for all market participants. In this paper, the profit for each market participants has been considered in a neutral market with double auction. Based on these profit functions, an optimal quantity of bids and offers is determined for buyers and sellers. Then a model based on non-cooperative game theory is proposed that leads to recognize optimal strategies for game players that include buyers, sellers and the market owner in complete information situation. Using the optimal strategies will satisfy the game participants and maximize their benefits. With the help of the obtained information, Market owner can increase benefits and support the markets profitability in a long time. At the end of the paper a numerical simulation is presented to show the role of game theory in the model and how game theory helps.

Keywords: Business2Business, E-Business, Buyer-Seller Relationship, Game Theory.

چکیده: بازارهای الکترونیکی B2B تسهیل کننده تراکنش های بین کسب و کارها هستند. متأسفانه تعداد زیادی از این بازارها نتوانسته اند منافع پیش بینی شده را فراهم کنند. در متن حاضر سود مورد انتظار هر یک از شرکای بازار در یک بازار خنثی و مزایده دوطرفه در نظر گرفته شده است. بر پایه توابع سود، مقدار بهینه ای برای سفارش خریدار و فروشنده تعیین شده است. سپس مدلی بر پایه تئوری بازی غیر همکاری در حالت اطلاعات کامل پیشنهاد شد. که این مدل منجر به شناسایی استراتژی های بهینه برای بازیکنان شامل، خریداران، فروشندگان و مالک بازار گردید. استفاده از استراتژی های بهینه موجب افزایش سود شرکای بازار و نهایتاً رضایتمندی آنها می شود. به کمک اطلاعات به دست آمده مالک بازار می تواند منفعت بازار را افزایش داده و سودآوری آن را در طولانی مدت تضمین کند. در بخش آخر متن، برای نشان دادن نقش تئوری بازی ها مثالی عددی نمایش داده شده است.

کلمات کلیدی: B2B، کسب و کار الکترونیکی، رابطه فروشنده-خریدار، تئوری بازی ها.

1- Introduction

B2B electronic market is an online market that facilitates the transactions between businesses to exchange their fabric and products. A number of factors are identified which impact on the buyer-supplier interface in B2B commerce by Humphreys et al [7]. The success of Business to Business (B2B) e-commerce accurately predicted by Wang et al. for small and medium enterprises [17]. They suggested that the implementation of B2B e-commerce is time consuming and the long-term impact on an organization may be unclear for some time. Unfortunately, the adoption of B2B electronic marketplaces was slower than predictions. Some of B2B electronic market failure reasons are:

- Business models of most B2B exchanges were not well developed [18].

- There was a lack of standards to promote integration across company sites, distributor sites, and industry sponsored exchanges with back end systems [14].
- Companies decide to turn inward and focus on private exchanges that linked them with key partners and suppliers [13].
- Suppliers are unwilling or unable to participate [14].

In the research of Wang et al. a mathematical model is introduced that compared the condition of buyer and seller in both online (electronic) market and offline (traditional) market. This model determines the situation that buyer and seller will satisfy in the electronic market [16]. In spite of this mathematical model, they considered the integrated and decentralized supply chain. Berg et al., used a probabilistic fuzzy modeling for financial markets analysis [2].

Against these mathematical models, some game theoretic models are proposed for market analysis. Gan et al. showed an auction game model for pool-based electricity market [5]. They claimed that the introduced concept of quasi-equilibrium offers an alternative for market studies. Kang et al. presented a non-cooperative game theory concept in single auction power pool to win the auction [8]. They considered a game between two suppliers, rather than supplier and buyer. They used marginal cost-a set of costs like maintenance, operational and upgrade costs- and the price per unit of power for each power generators to determine the result of auction. Also they performed demand forecasting. Their methodology gives optimal bidding strategies for competitive power suppliers.

A game theory simulator for assessing the performance of competitive electricity markets proposed by Bompard[3]. The simulator is run for one year. According to their research, game theory can be applied to simulate realistic market framework. Behavioral procedures simulated by Menniti et al. for electricity market with Genetic Algorithm (GA) as an evolutionary game [11]. They used GA to forecast the electricity price and how the competition can influence it over a long period. Lise et al. investigated market power and the environmental effects of the Northwestern European electricity market with a game theoretic model [10]. Rosenthal used game theory to illustrate the transfer pricing in a vertically integrated supply chain [15].

A computer-based learning environment is introduced as a microworld for understanding risks in a deregulated industry [4]. Ganeshan et al. proposed an optimal procurement portfolio in B2B market to help a procurement manager [6]. Munksgaard et al. focused on tariffs and investments determination to grow the Danish liberalised power market [12]. A n-person noncooperative bargaining game suggested by Kim et al. that the game leads to a nonlinear programming function, but unfortunately the model may not be directly apply to some cases [9].

In this paper the expected profit of each market participants has produced, and then using the game theory, the optimal mixed strategies for owner is presented. Owner can assign offers and bids together in a more profitable manner. The rest of the paper structured as follow. In section 2, the trading model in double auction and the functions of expected profit for buyers, sellers and the market owner is proposed. In section3, the trade's game and the

mixed strategies in a game theory is described. Then a numeric example to explain the model is provided in section 4. Finally, the conclusion and a short discussion are given in section 5.

2- The proposed trading model in double auction

In a double auction, sellers and buyers make offer and demand bids and send them to the auctioneer (market owner) as depicted in figure 1.

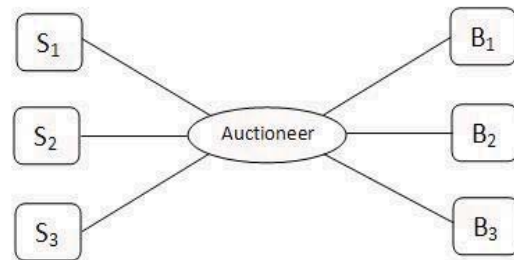


Figure 1: Buying and selling in double auction

The prices suggested by buyer is P_b and by seller is P_s . Corresponding value for the good suggested by buyer is V_b . Consider that $P_b < V_b$ because buyer do not bid a price more than attributing value. The market owner receives the offers and bids so makes decision to how to assign the offers and bids together. There are two conditions:

$$\begin{cases} \text{If } P_{si} \leq P_{bj} & \text{Then The trade is possible} \\ \text{If } P_{si} > P_{bj} & \text{Then Nothing} \end{cases}$$

2-1- Notations

We defined the following quantities:

P_s : proposed price by buyer

P_b : Proposed price by seller

V_b : Corresponding value suggested by buyer

α : constant cost charged by the auctioneer to the seller

β : constant cost charged by the auctioneer to the buyer

x : nonnegative continuous random variable for demand

$f(x)$: Probability density function for random variable x

$F(x)$: Cumulative distribution function for random variable x , $F(x)$ is continuous, differentiable, invertible and strictly increasing

C : Seller's per unit production cost

2-2- Assumptions

We will use the following assumptions the rest of the papers:

- $P_b \leq V_b$. This assumption used to avoid trivial problems.
- P_s is the trade's price. $P_b - P_s$ charged by auctioneer.

2-3-The proposed model

The seller, buyer and owner (auctioneer)'s expected profit can be expressed as table 1 when sellers and buyers only pay a constant cost. We assumed different constant costs for buyer and sellers just for increasing the flexibility of the model, absolutely these can consider as equal.

Table 1: Expected profit function for sellers, buyers and the market owner

	If $P_s \leq P_b$	If $P_s > P_b$
Seller's payoff	$\int_0^Q P_s x f(x) dx + \int_Q^\infty P_s Q f(x) dx - CQ$ $-\alpha$	$-\alpha$
Buyer's payoff	$\int_0^Q (V_b - P_b) x f(x) dx + \int_Q^\infty (V_b - P_b) Q f(x) dx$ $-\beta$	$-\beta$
Owner's payoff	$(P_b - P_s)Q + \alpha + \beta$	$\alpha + \beta$

$-\alpha$ shows that if $P_s > P_b$ seller has lost α and same for buyer, if $P_s > P_b$ buyer has lost β .

Seller payoff (Q) =

$$\int_0^Q P_s x f(x) dx + \int_Q^\infty P_s Q f(x) dx - CQ - \alpha$$

This function is strictly concave. So there is a unique optimal quantity that maximizes the seller's payoff. This optimal quantity can calculate as:

$$Seller\ payoff'(Q) = -F(Q) * P_s + P_s - C = 0 \tag{1}$$

So,

$$Q_s^* = F^{-1}\left(\frac{P_s - C}{P_s}\right) \tag{2}$$

$$buyer\ payoff(Q) = \int_0^Q (V_b - P_b) x f(x) dx + \int_Q^\infty (V_b - P_b) Q f(x) dx - \beta$$

and in the same manner the optimal quantity for buyer is:

$$Q_b^* = F^{-1}\left(\frac{V_b - P_b}{V_b - P_b}\right) = F^{-1} \tag{3}$$

Seller knows about α , C and his proposed price, P_s , so he can calculate the optimal quantity of order. For buyer, $F(Q) = 1$, it means, there is a special Q that satisfy the buyer and this point is a $q \leq Q^*$ and $\forall Q, Q = Q^* \dots \infty F(Q) = 1$. If $P_s \leq P_b$ this model will satisfy the buyers but if $P_s > P_b$ they will lose β .

Also, Q_b^* and Q_s^* are not depend on α and β . It shows that quantity of order and bid of market participants relates to their own information like production cost or corresponding value for product.

3- The trade's game

Game theory is a rich area of mathematics for economics, politics, finance, military science, and so on. Suppose that there are n seller and m buyer. Owner will construct a matrix game to show the game situation between each seller and buyer. Every entry in the matrix poses the three values, market participants expected profits, seller's payoff (S_{ij}), buyer's payoff (B_{ij}), and owner's payoff (O_{ij}): $a_{ij} = (S_{ij}, B_{ij}, O_{ij})$.

$$A = \begin{pmatrix} (S_{11}, B_{11}, O_{11}) & (S_{12}, B_{12}, O_{12}) & \dots & (S_{1m}, B_{1m}, O_{1m}) \\ (S_{21}, B_{21}, O_{21}) & \vdots & \vdots & (S_{2m}, B_{2m}, O_{2m}) \\ \vdots & \vdots & \vdots & \vdots \\ (S_{n1}, B_{n1}, O_{n1}) & \vdots & \vdots & (S_{nm}, B_{nm}, O_{nm}) \end{pmatrix}$$

$$S_{ij} = \begin{cases} \text{If } P_{si} \leq P_{bj} \text{ Then } P_{si} Q_{bj}^* - C Q_{bj}^* - \alpha \\ \text{If } P_{si} > P_{bj} \text{ Then } -\alpha \end{cases}$$

$$B_{ij} = \begin{cases} \text{If } P_{si} \leq P_{bj} \text{ Then } (V_{bj} - P_{bj}) Q_{bj}^* - \beta \\ \text{If } P_{si} > P_{bj} \text{ Then } -\beta \end{cases}$$

$$O_{ij} = \begin{cases} \text{If } P_{si} \leq P_{bj} \text{ Then } (P_{bj} - P_{si}) Q_{bj}^* + \alpha + \beta \\ \text{If } P_{si} > P_{bj} \text{ Then } \alpha + \beta \end{cases}$$

a_{ij} appears the game's payoff when i^{th} seller is trading with j^{th} buyer. Consequently, we will have a 5-dimensional game matrix that the dimensions are:

- 1- number of sellers
- 2- number of buyers
- 3- number of each sellers strategies
- 4- number of each buyers strategies
- 5- quantity (Q)

A mixed strategy is vector $X = (x_1, \dots, x_n)$ for seller selection and $Y = (y_1, \dots, y_m)$ for buyer selection that x_i is the probability of i^{th} seller will trade and y_j is the probability of j^{th} buyer participation. And $\sum_{i=1}^n x_i = 1$ and $\sum_{j=1}^m y_j = 1$. Expected payoff in mixed strategy generates as below:

$$E(X, Y) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} \text{prob}(S \text{ uses } i) \text{prob}(B \text{ uses } j) = \sum_{i=1}^n \sum_{j=1}^m x_i a_{ij} y_j = XAY^T = [x_1 \dots x_n] \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \tag{4}$$

According to [1], we can calculate the v, X, Y , (v is the value of the game) and consequently a pair (x^*, y^*) that this is a saddle point for the game.

4- The numeric example

Suppose that there are 4 options for seller and 4 options for buyer. For example each seller can sell her products in a different delivery times or types (Post the product Online, by train, by plain, before the payment, after money received, and etc). And each buyer has different payment methods and policies. a_{ij} in matrix A shows the game's payoff in a situation that S is trading with his i^{th} strategy and B is trading with her j^{th} strategy on Q number of product. So we fixed 3 dimensions because we supposed that one of the buyers is trading with one of the sellers for specified number of quantity. Matrix A can be considered from the all market participants viewpoints to help the analysis in complete information. a_{ij} will be calculated from proposed model for buyer and seller for each strategy.

We generated the hypothesis matrix from a seller's standpoint to illustrate the model like this:

$$A = \begin{pmatrix} 5 & 0 & -2 & 6 \\ 2 & 3 & 2 & 4 \\ -3 & 1 & 1 & 1 \\ 0 & 0 & -4 & 3 \end{pmatrix}$$

To find the Y and a saddle point, Theorem 1 part (a) and (b) is been used [1].

Theorem 1:

- Part (a): If $E(i, Y^*) \leq w \leq E(X^*, j), i = 1, \dots, n, j = 1, \dots, m$, then $w = v(A)$ and (x^*, y^*) is a saddle point for the game.
- Part (b): A strategy X^* for player 1 is optimal if and only if $v(A) = \min_{1 \leq j \leq m} E(X^*, j)$. A strategy Y^* for player 2 is optimal if and only if $v(A) = \max_{1 \leq i \leq n} E(i, Y^*)$

So we have:

$$\begin{aligned} E(1, Y^*) &= v \rightarrow 5y_1 - 2y_2 + 6y_4 = v \\ E(2, Y^*) &= v \rightarrow 2y_1 + 3y_2 + 4y_3 = v \\ E(3, Y^*) &= v \rightarrow -3y_1 + y_2 + y_3 + y_4 = v \\ E(4, Y^*) &= v \rightarrow -4y_3 + 3y_4 = v \end{aligned}$$

and

$$y_1 + y_2 + y_3 + y_4 = 1$$

Then,

$$Y = \left(\frac{-46}{49}, \frac{11}{49}, \frac{22}{49}, \frac{62}{49} \right)$$

which is not a strategy because $\exists y_j < 0$. This means that our assumption about the existence of an optimal strategy with $y_j > 0, j = 1, 2, 3, 4$ must be wrong. We know that $\min E(2, y) = 2$. For an optimal strategy for buyers, as well as $E(1, y) < 2, E(3, y) < 2$, and $E(4, y) < 2$. So that $y_1 + y_2 + y_3 + y_4 = 1$. The plot of these inequalities showed in figure 2.

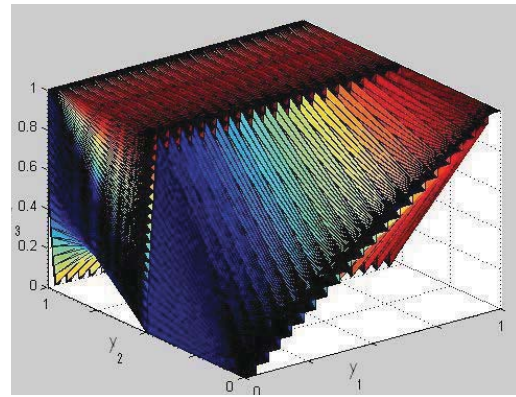


Figure 2: Optimal strategy set for Y

There are lots of points which work. In particular $Y = (0, 0.1, 0.6, 0.3)$ will give an optimal strategy for buyers from sellers standpoint (with seller's payoff function). So the market owner can calculate the optimal mixed strategies for sellers from seller's standpoint.

$$\begin{aligned} E(X^*, 1) &= v \rightarrow 5x_1 + 2x_2 - 3x_3 = 2 \\ E(X^*, 2) &= v \rightarrow 3x_2 + x_3 = 2 \\ E(X^*, 3) &= v \rightarrow -2x_1 + 2x_2 + x_3 - 4x_4 = 2 \\ E(X^*, 4) &= v \rightarrow 6x_1 + 4x_2 + x_3 + 3x_4 = 2 \end{aligned}$$

$$\text{and } x_1 + x_2 + x_3 + x_4 = 1$$

So,

$$X = \left(\frac{100}{67}, \frac{52}{67}, \frac{-22}{67}, \frac{-63}{67} \right)$$

which is not a strategy because $\exists x_i < 0$. It is necessary to look for x_1, x_2, x_3, x_4 to find an optimal strategy for S.

$$\begin{aligned} E(X^*, 1) &> 2 \rightarrow 5x_1 + 2x_2 - 3x_3 > 2 \\ E(X^*, 2) &> 2 \rightarrow 3x_2 + x_3 > 2 \\ E(X^*, 3) &> 2 \rightarrow -2x_1 + 2x_2 + x_3 - 4x_4 > 2 \\ E(X^*, 4) &> v \rightarrow 6x_1 + 4x_2 + x_3 + 3x_4 > 2 \end{aligned}$$

So replaced $x_4 = 1 - x_1 - x_2 - x_3$ and then get a graph of the region of points satisfying all the inequalities in (x_1, x_2, x_3) space.

Inequalities:

$$\begin{aligned} 5x_1 + 2x_2 - 3x_3 &> 2 \\ 3x_2 + x_3 &> 2 \\ -2x_1 + 2x_2 + x_3 - 4x_4 &> 2 \\ 6x_1 + 4x_2 + x_3 + 3x_4 &> 2 \end{aligned}$$

Also owner can construct the same matrix from his/her self standpoint or buyers standpoint (using buyers payoff function) to obtain the mixed strategies for all market participants. Thus, the market owner has clear information about market and he/she can make the best assignments to satisfy maximum number of sellers and buyers and consequently increases the market profitability. It

leads to encourage the game participants to stay in the market.

5- Conclusion

This paper describes the expected profit in an electronic market that uses the double auction. The proposed model has a series of separated expected profit functions for sellers, buyers and market owner (auctioneer). The model has been used in a game theory as a payoff function so that the game matrix is a 5-dimension matrix. A numerical simulation is represented to convey the concept and application of the model when a seller is trading with a buyer on a specific quantity. This model works in a complete information situation. We can expand this model to an incomplete information condition in the future.

References

- [1] Barron E. N., *Game theory: an introduction*, published by John Wiley and Sons, Inc., (2008).
- [2] Berg J., Kaymak U., Berg W. M., *Financial markets analysis by using a probabilistic fuzzy modelling approach*, International Journal of Approximate Reasoning 35, (2004), 291-305.
- [3] Bompard E., Carpaneto E., Ciwei G., Napoli R., Benini M., Gallanti M., Migliavacca G., *A game theory simulator for the performances of competitive electricity markets*, Electronic Power Systems Research 78, (2008), 217-227.
- [4] Dyer I., Larsen E., Franco C. J., *Games for electricity traders: understanding risk in a deregulated industry Energy policy*, (2008), doi:10.1016/j.enpol.2008.09.075
- [5] Gan D., Wang J., Bourcier D. V., *An auction game model for pool-based electricity markets*, Electrical Power and Energy Systems 27, (2005), 480-487.
- [6] Ganeshan R., Boone T., Aggarwal P., *Optimal procurement portfolios when using B2Bs: A model and analysis*, International Journal of Production Economics, (2008), doi: 10.1016/j.ijpe.2008.08.041.
- [7] Humphreys P., McIvor R., Cadden T., *B2B commerce and its implications for the buyer-supplier interface*, 11, 2, (2006), 131 - 139.
- [8] Kang D. J., Kim B. H., Hur D., *Supplier bidding strategy based on non-cooperative game theory concepts in single auction power pools*, Electronic Power System Research 77, (2007), 630-636.
- [9] Kim T., Jeon Y., *Stationary perfect equilibria of an n-person noncooperative bargaining game and cooperative solution concepts*, European Journal of Operational Research 194 (2009) 922-932.
- [10] Lise W., Linderhof V., Kuik O., Kemfert C., Ostling R., Heinzow T., *A game theoretic model of the Northwestern european electricity market-market power and the environment*, Energy Policy 34, (2006), 2123-2136.
- [11] Menniti D., Pinnarell A., Sorrentino N., *Simulation of producer behaviour in the electricity market by evolutionary games*, Electronic Power Systems Research 78, (2008), 475-483.
- [12] Munksgaard J., Morthorst P. E., *Wind power in the Danish liberalised power market-Policimeasures, price impact and investor incentives*, Energy Policy, 36, (2008), 3940-3947.
- [13] Murphy J., and Barling B., *Sell-Side product content: the key to supplier empowerment*, AMR research, Inc., (July 2001)
- [14] Pang A., and Villars R., *Second-Generation B2B e-commerce*, IDC White Paper, (August 2001).
- [15] Rosenthal E. C., *A game-theoretic approach to transfer pricing in a vertically integrated supply chain*, International Journal of Production Economics 115, (2008), 542-552.
- [16] Wang C. X., Benaroch M., *Supply chain coordination in buyer centric B2B electronic markets*, International Journal of Production Economics 92, (2004), 113-124.
- [17] Wang T. C., Lin Y. L., *Accurately predicting the success of B2B e-commerce in small and medium enterprises*, Expert Systems with Applications, to be published, (2008).
- [18] Wise R., Morrison D., *Beyond the exchange the future of B2B*, Harvard Business Review, (November-December 2000).



Journal of Control

A Publication of Iranian Society of Instrumentation and Control Engineers, Vol. 2, No. 2, Winter 2009.

Publisher: Iranian Society of Instrumentation and Control Engineers

Managing Director: Prof. Iraj Goodarznia

Editor-in-Chief: Prof. Ali Khaki-Sedigh

Tel: 84062317

Email: sedigh@kntu.ac.ir

Assistant Editor: Dr. Hamid Khaloozadeh, Dr. Alireza Fatehi

Executive Director: Dr. Hamid Khaloozadeh

Editorial Board:

Prof. A. Khaki-Sedigh, Dr. H. Khaloozadeh (Associate Prof.), Prof. I. Goodarznia, Prof. P. Jabedar-Maralani, Prof. A. Ghafari, Dr. H. Hassibi (Assistant Prof.), Dr. M.R. Jahed-Motlagh (Associate Prof.), Dr. K. Badie, Prof. R. Asgharian, Dr. H.R. Momeni (Associate Prof.), Prof. A. Vahidian-Kamyad, Prof. S. Khanmohammadi, Prof. S.K. Nikraves, Prof. M. Shafiee, Dr. A. Fatehi (Assistant Prof.), Dr. B. Labibi (Assistant Prof.), Prof. B. Moshiri, Prof. S. Katebi, Dr. A. Maghsoudi-Pour, K. Falamaki (M.Sc.), M. Baradaran-Mozafari (M.Sc.), Gh. A. Ramazani (M.Sc.).

Advisory Board:

Dr. H.R. Momeni, Dr. N. Abedi, Prof. A. Ghaffari, Dr. A. A. Gharehveisi, Dr. M. Tavakoli-Bina, Dr. H.R. Taghirad, Dr. M. Bathaei, Dr. M.T. Hamidi-Beheshti, Prof. B. Moshiri, Prof. M. Shafiee, Prof. R. Asgharian, Prof. A. Khaki-Sedigh, Dr. B. Moaveni, Dr. R. Kazemi, Dr. S.A. Mousavian, Dr. A.H. Markazi-Davaei, Prof. M. Haeri, Dr. A.R. Khalili-Tehrani, Prof. H. Seifi, Dr. A. Kazemi, Dr. H. Khaloozadeh, Dr. M. Aliari-Shourehdeli, Dr. A. Fatehi, Dr. M.R. Akbarzadeh-Toutounchi, Dr. Mirabedini, Dr. H. Pedram, Dr. A. Harounabadi, Prof. A. Vahidian-Kamyad, Dr. M. Arvan, Dr. J. Heirani-Nobari, Dr. B. Labibi, Prof. F. Hossein-Babaei, Dr. K. Aghaei.

The ISICE Board of Director:

A. Sheri-Moghadam, Dr. K. Masroori, Dr. H.R. Momeni, Prof. B. Moshiri, Dr. F. Jafar-Kazemi, Dr. H. Khaloozadeh, A. Rastegari

Address: Unit 241, 2nd floor, No.27, Mousavi Ave. Ferdowsi Sq. Enghelab St. Tehran, Iran.

P.O. Box: 15815-3595

Tel: (+9821) 88813002

Fax: (+9821) 88324979

<http://www.isice.ir>



Journal of Control

A Publication of Iranian Society of Instrument and Control Engineers
Vol. 2, No. 2, Winter 2009

Persian Part

- ۱ ارائه روش مقاوم جهت ترکیب داده در سیستم تلفیقی و مقایسه آن با فیلتر کالمن توسعه یافته
در شرایط بحرانی GPS/SDINS
علی اسدیان، بهزاد مشیری، علی خاکی صدیق
- ۱۲ طراحی مطمئن کنترل کننده H_{∞} برای سیستم های خطی تأخیردار با وجود نامعینی در تمام
ماتریس های سیستم
علی ابوطالبی، سعید حسین نیا، فرید شیخ الاسلام
- ۲۰ ارائه روشی جدید در طراحی کنترلگر تطبیقی مقاوم برای سیستم های تأخیردار
سید کمال حسینی ثانی، حمیدرضا مؤمنی
- ۲۸ تأثیر بردار برتر در دینامیک فرآیند های QRDS و کاربرد آن در کنترل فرآیندهای دارای
تأخیر زمانی
منصور شیروانی، منصوره اسماعیلی

English Part

- The Application of Non-Cooperative Game Theory in B2B Electronic Market 1
Sahar Sohrabi, Mohammad Jafar Tarokh
- New Steady State Kalman Filter for Tracking High Maneuvering Targets 8
Seyyed-Amin Hosseini, Mohammad Farrokhi
- A Behavioral Approach to Estimate Area of Expertise for Cooperative Q-Learning: A Self Referenced Strategy 25
Ahmad Imanipour, Majid Nili Ahmadabadi, Babak N. Araabi, Masoud Asadpour,
Roland Siegwar, Aram Zaerpour