



توسعه یک مدل کنترل بهینه در تصمیم‌گیری بنگاه‌های انحصار دوقطبی با استفاده از تعادل نش و استکلبرگ

سید جلال ضیایی (نویسنده مسئول)

گروه مهندسی صنایع، دانشکده مهندسی صنایع، دانشگاه صنعتی شریف، تهران، ایران

Email: ziayei@yahoo.com

رسول حجی

استاد دانشکده مهندسی صنایع، دانشگاه صنعتی شریف، تهران، ایران

تاریخ دریافت: ۱۴۰۳/۱۱/۱۰ * تاریخ پذیرش: ۱۴۰۴/۰۸/۱۰

فصلنامه مدیریت راهبردی در سیستم‌های صنعتی (مدیریت صنعتی سابق)
دانشگاه آزاد اسلامی واحد سنجند
سال بیستم / شماره ۷۳ / پاییز ۱۴۰۴

چکیده

در محیط پر از چالش کسب و کار امروزی، استفاده از روش‌های مناسب تصمیم‌گیری در شرایط تعارض برای اتخاذ تصمیم مناسب با در نظر گرفتن تصمیم‌های رقبا بیش از پیش احساس می‌شود. در دنیای واقعی تصمیمات بنگاه‌ها علاوه بر تصمیم بنگاه رقیب به دوره تصمیم‌گیری نیز وابسته است به این‌صورت که ممکن است در دوره تحت بررسی عواملی از قبیل اشباع شدن بازار، تغییر ذائقه مشتریان و یا حتی فصلی بودن محصول در استراتژی انتخابی بنگاه‌ها مؤثر واقع گردد. در این مقاله، ابتدا مسائل کنترل بهینه گسسته معرفی و فرمول‌بندی مربوط به آنها ارائه گردیده است. فرمول‌بندی مربوطه برای مدل کنترل، به همراه تکنیک‌های توسعه داده شده برای برنامه‌ریزی پویا، امکان حل کلاس بزرگی از مسائل تصمیم را فراهم می‌آورد. مدل ارائه شده به گونه‌ای توسعه داده شده‌است تا حالت خاصی از بازی‌های دونفره‌ای که در آنها زمان به عنوان یک پارامتر مستقل در پیامد بازیکنان مؤثر است را حل نماید. برای تشریح مدل مربوطه حالات خاصی از مدل انحصار دوقطبی بررسی می‌گردد که در آن دو بنگاه تولیدی که محصولات مشابهی تولید می‌کنند در پی این موضوع هستند که چه میزان کالا را در هر دوره، با توجه به ظرفیت تولید Π دوره‌ای، به بازار عرضه نمایند تا سودشان در مجموع Π دوره ماکزیمم گردد. بدین‌منظور تکنیکی ترکیبی ارائه شده است که از مفاهیم موجود در نظریه بازی‌ها یعنی تعادل نش و استکلبرگ در مدل کنترل بهینه بهره می‌برد. جواب‌های بدست آمده از مدل تلفیقی، کارایی تکنیک پیشنهادی را اثبات می‌کند. برای تأیید اثربخشی تکنیک ترکیبی، دو مثال عددی ارائه و روش‌شناسی برای حل آنها بکار گرفته شده است. مثال‌های عددی و تحلیل نتایج، بینش مدیریتی مفید جهت انتخاب بهترین استراتژی‌ها برای بنگاه‌ها در محیط‌های مختلف بازار و عملیاتی را فراهم می‌آورد.

کلمات کلیدی: نظریه بازی‌ها، تعادل نش، تعادل استکلبرگ، کنترل بهینه گسسته، برنامه‌ریزی پویا، مدل انحصار دوقطبی.

۱- مقدمه

در هنگام مواجهه با بسیاری از مسائل اقتصادی، اجتماعی، سیاسی و نظامی، اغلب ناگزیر به تجزیه تحلیل موقعیت‌هایی هستیم که در آنها دو یا چند رقیب اهداف متعارضی را دنبال می‌کنند و نتیجه تصمیم اتخاذی آنها بستگی به خط مشی انتخابی رقبا دارد، چنین موقعیت‌هایی را موقعیت‌های تعارض‌آمیز^۱ می‌نامیم. در محیط پر از چالش کسب‌وکار امروزی نمی‌توان تنها با استفاده از روش‌های تصمیم‌گیری سنتی برای مواجهه با رقبا تصمیم‌های صحیحی اتخاذ نمود؛ بنابراین استفاده از روش‌های مناسب تصمیم‌گیری در شرایط تعارض برای اتخاذ تصمیم مناسب با در نظر گرفتن تصمیم‌های رقبا بیش از پیش احساس می‌شود. از طرفی دیگر، عواملی چون پیشرفت فناوری، نظریه‌های نوین اقتصادی، بازاریابی و نوسانات سیاسی، بر پویا بودن این تصمیم‌ها تأثیر گذاشته است؛ به این معنی که لازمه اتخاذ تصمیم مناسب برای شرکت‌ها، توجه به تصمیم رقبا و تغییرات آن‌ها در طول زمان است. در واقع مسأله عبارت است از بهینه‌سازی متغیرهای توانمندساز یک کسب‌وکار مانند میزان تولید و عرضه، با در نظر گرفتن تعاملات و موقعیت‌های تعارض‌آمیزی که شرکت‌ها با رقبایشان درگیر هستند.

در علم اقتصاد، ساختارهای متفاوتی برای بازار تعریف می‌شود که موضوع رقابت در آن‌ها متفاوت است (ساختار بازار معرف خصوصیات سازمانی بازار است و معمولاً برحسب سطح تمرکز، تفاوت کالا و شدت موانع ورود به آن تعریف می‌شود. ساختار بازار براساس تعداد شرکت‌های حاضر در بازار برای فروش یک محصول خاص و یا ارائه خدمتی مشخص به بازار رقابت کامل (تعدد فروشنده و خریدار)، بازار انحصار کامل یا منوپلی^۲ (یک فروشنده و تعداد زیادی خریدار)، بازار انحصار چندجانبه یا آلیگوپلی^۳ (تعداد محدودی فروشنده و تعدد خریدار) تقسیم می‌شود. اگر در بازار انحصار چندجانبه تعداد فروشندگان دو باشد، بازار انحصار دوجانبه یا دوپلی^۴ نامیده می‌شود (Raoufinia et al., 2016). در ساختار الیگوپلی شرکت‌ها معمولاً با نوعی تردید و دوراهی روبرو هستند. شرکت‌ها می‌توانند راهبردهای سوددهی صنعت را بدون در نظر گرفتن راهبردهای رقبا دنبال کنند و خطر واکنش تلافی‌جویانه رقبا و افزایش رقابت و مبارزه در درون صنعت را بپذیرند و یا اینکه راهبردهای خود را براساس تحرکات رقبا برای کسب موفقیت در بازار طرح‌ریزی کنند (Basiri et al., 2016). تصمیم‌گیری در مورد سیاست‌های تولیدی و قیمتی هر عرضه کننده در بازار انحصاری چندجانبه با در نظر گرفتن این امر که عکس‌العمل رقبا در قبال هر سیاست چه خواهد بود موضوعی است که تولیدکنندگان در این بازارها با آن روبرو هستند؛ زیرا هر تغییری در میزان قیمت یا مقدار تولید یک بنگاه اقتصادی، فروش و سود رقبا را تحت تأثیر قرار می‌دهد (Shahikitash & Mohammadzadeh, 2014). لذا مسأله تحلیل بازارهای انحصاری چندجانبه و تدوین راهبردهای مناسب برای فروشندگان با در نظر گرفتن عکس‌العمل رقبا، ضمن آنکه پیچیدگی خاصی نسبت به تحلیل سایر ساختارهای بازار دارد، دارای اهمیت ویژه‌ای است (Raoufinia et al., 2016).

رویکرد پیشنهادی در این پژوهش برای حل مسأله یادشده و کمک به شرکت‌ها برای تحلیل موقعیت شرکت و رقبا و بهینه‌سازی متغیرهای تصمیم برای توانمندسازی در بازار رقابتی، رویکرد نظریه بازی‌ها از نوع پویا وابسته به پارامتر زمان است. نظریه بازی‌ها در تلاش است توسط ریاضیات، رفتار را در شرایط راهبردی یا بازی که در آن‌ها موفقیت فرد در انتخاب کردن، وابسته به انتخاب دیگران است بدست آورد. وابستگی یعنی اینکه هر بازیکن از آنچه که دیگران در بازی انجام می‌دهند، تأثیر می‌پذیرد و رفتار بازیکنان نیز بر دیگران تأثیر می‌گذارد. خروجی بازی وابسته به تصمیم‌های همه افراد است و هیچ فردی بر آنچه اتفاق می‌افتد، کنترل ندارد. افراد این وابستگی‌های متقابل را می‌دانند و آن‌ها را در تصمیم‌های خود لحاظ می‌کنند (Milan, 1996). یک بازی شامل مجموعه‌ای از بازیکنان، مجموعه‌ای از حرکات یا راهبردها و نتیجه مشخصی برای هر ترکیب از راهبردها است. پیروزی در هر بازی تنها تابع یاری شانس نیست. بلکه اصول و قوانین ویژه خود را دارد و هر بازیکن طی بازی سعی می‌کند با بکارگیری آن اصول خود را به برد نزدیک کند. نظریه بازی تلاش می‌کند تا رفتار ریاضی حاکم بر یک موقعیت راهبردی (تضاد منافع) را مدلسازی کند. این موقعیت زمانی پدید می‌آید که موفقیت یک فرد وابسته به راهبردهایی است که دیگران انتخاب می‌-

¹ Conflict situation² Monopoly³ Oligopoly⁴ Duopoly

کنند. هدف نهایی این دانش، یافتن راهبرد بهینه برای بازیکنان است (Raoufinia et al., 2016). هنگامی که رویکرد نظریه بازی به عنوان راه حلی برای مشکلات مربوط به تصمیم‌گیری به کار می‌رود، بررسی اینکه کدامیک از مفاهیم و روشهای محاسباتی منطبق با آن باید مورد استفاده قرار گیرد، حائز اهمیت بوده و برای به اجرا درآوردن نتایج بررسی ضروری است. اگر پارامترها و مدل‌های ریاضی برای حل مشکلات مربوط به تصمیم‌گیری بدون توجه به عدم قطعیتی که احتمالاً در سیستمهای رقابتی روی خواهد داد تعیین شوند، نتایج تحلیل و حل مشکلات مربوط به تصمیم‌گیری همواره برای مشکلات واقعی مناسب نخواهد بود. زمان یا دوره انجام بازی به عنوان یک پارامتر مستقل در ایجاد این نااطمینانی و عدم قطعیت، تأثیرگذار بوده و باعث ایجاد ابهام در نظرات و احکام و نهایتاً تصمیم مناسب بازیکنان می‌گردد. مدل ارائه شده در این مقاله به دنبال دستیابی به تصمیم مناسب در شرایط ذکر شده است.

در این پژوهش بازی‌های دو نفره‌ای مورد بررسی قرار می‌گیرد که پیامدهای موجود در ماتریس پرداخت آنها وابسته به پارامتر زمان بوده و تابعی گسسته از زمان است. هدف پژوهش حاضر، برآورد ماتریس پرداخت بازی نبوده و بازی با اطلاعات کامل فرض شده است. یعنی هر بازیکن از مجموعه استراتژی‌های حریف و پیامدی که از ترکیبات مختلف استراتژی‌ها حاصل می‌گردد مطلع است. لذا ابتدا مسائل کنترل بهینه گسسته معرفی و فرمول‌بندی مربوط به آنها آورده شده سپس یک الگوریتم برای حل آنها ارائه گردیده است. سپس بازی‌های ماتریسی^۱ و دوماتریسی^۲ با پیامدهای وابسته به زمان، مدلسازی و با استفاده از الگوریتم ارائه شده برای حل مسائل کنترل بهینه گسسته حل می‌گردد. در ادامه رویکرد ارائه شده به بازی‌های با مجموعه استراتژی‌های پیوسته^۳ تعمیم و استراتژی‌های بهینه مدل کورنو n دوره‌ای با در نظر گرفتن نرخ تنزیل محاسبه شده است. برتری جوابهای حاصله از تکنیک ترکیبی نظریه بازی و کنترل بهینه، نسبت به جوابهای حاصل از تحلیل صرف نظریه بازی، کارایی تکنیک معرفی شده را مشخص می‌سازد.

۲- روش‌شناسی پژوهش

مطالعه حاضر از منظر نوع هدف تحقیق کاربردی توسعه یافته است و با استفاده از مفاهیم موجود در نظریه بازی‌ها یعنی تعادل نش و استکلبرگ مدل‌های انحصار دوقطبی را با استفاده از مدل‌های موجود در نظریه کنترل بهینه گسسته، مدل‌سازی و در پی آن نقاط تعادل بازی را بدست می‌آورد. به عبارتی دیگر، تکنیک تلفیقی نظریه بازی و کنترل بهینه ارائه شده در این پژوهش، راهبردهای بهینه را به بازیکنان ارائه می‌نماید. درواقع تعیین مقدار بهینه متغیرهای تصمیم، یعنی میزان عرضه کالا در بازارهای انحصار دوجانبه یا دوقطبی با در نظر گرفتن تصمیم‌های رقبا و پویایی وابسته به عامل زمان، موضوع این پژوهش است. لذا در این مقاله کلاس بزرگی از مسائل کنترل بهینه مطرح و یک الگوریتم برای حل آنها ارائه می‌گردد. رویکرد مطرح‌شده به بازی‌های دونفره با مجموع ثابت، غیرثابت و پیوسته تعمیم و برای نشان دادن کاربردهای عملی نتایج، بر روی حالات خاصی از مدل‌های انحصار دوقطبی پیاده‌سازی می‌گردد. برای تأیید اثربخشی تکنیک تلفیقی پیشنهادی، دو مثال عددی ارائه و روش‌شناسی برای حل آنها بکارگرفته شده است.

الف) پیشینه تحقیق

جیمز والدگرو اولین کسی بود که در سال ۱۷۱۳ با ارائه راه حل مینیماکس در مقاله‌ای نظریه بازیها را مورد بحث قرار داد. پس از او آکوستین کورنات در سال ۱۸۳۸ مقاله‌ای را تحت عنوان "تحقیقاتی در باب اصول ریاضی نظریه ثروت" منتشر کرد که در آن با ارائه راه حلی نقطه تعادل بازی انحصار دوقطبی را محاسبه نمود (Cournot, 1838). اگرچه تا آن زمان هنوز ایده تعادل نش^۴ توسط جان نش مطرح نشده بود اما بعدها معلوم شد که نقطه تعادل بدست آمده توسط کورنات، در واقع همان نقطه تعادل نش برای بازی انحصار دوقطبی است (Nash, 1950). ون نیومن در سال ۱۹۴۴ به همراه اقتصاددانی اتریشی به نام اسکار مورگنسترن کتاب "نظریه بازی‌ها و رفتار اقتصادی" را به رشته تحریر در آورد (Neumann & Morgenstern, 1944).

¹ Matrix game

² Bimatrix games

³ Game with Continuous Strategy Sets

⁴ Nash equilibrium

محققین مختلفی به بسط و توسعه تئوری بازی‌ها از طریق نظریه کنترل بهینه^۱، در قالب بازیهای دیفرانسیلی^۲ پرداخته اند. به طور مثال محمودی نیا و همکارانش در تحقیقی کاربرد این نوع بازیها را در تعیین سیاست‌های بهینه در تقابل استراتژیک بین سیاست‌گذاران پولی و مالی بیان کرده‌اند (Mahmudinia et al., 2017). تاتاری و نقیبی مدل کنترل بهینه توزیع شده سیستم‌های چند عاملی غیر خطی زمان پیوسته را توسط بازی‌های دیفرانسیلی ارائه کرده‌اند که در آنها دینامیک خطا و اندیس عملکرد هر عامل بستگی به عامل‌های همسایه‌اش دارد (Tatari & Naghibi, 2015). واموداکیس و لوئیس نیز الگوریتمی برای حل بازی‌های دیفرانسیلی مجموع غیر صفر به صورت زمان - حقیقی برای سیستم‌های زمان پیوسته غیر خطی و دینامیک کاملاً معین ارائه کرده‌اند (Vamvoudakis & Lewis, 2011) که ژانگ و همکارانش با ارائه روشی مشابه، پیچیدگی محاسباتی الگوریتم ارائه شده توسط آنها را کاهش داده اند (Zhang et al., 2013).

محققین مختلفی نیز از ترکیب نظریه بازی و برنامه ریزی پویا^۳ استفاده کرده اند. پیکل و لوزوانو یک مدل کنترل بهینه گسسته را به کمک تئوری بازیهای غیرمشارکتی روی حالت خاصی از شبکه‌های متناظر با گرافهای جهت دار k بخشی توسعه داده‌اند (Lozovanue & Pickl, 2006). بارکوئین و همکارانش در یک مطالعه موردی حالت خاصی از بازی های با پیامدهای پیوسته را با استفاده از برنامه‌ریزی پویا تحلیل نموده و مطلوبیت و مزیت جوابهای حاصله از ترکیب این دو ابزار قدرتمند بهینه-سازی یعنی برنامه ریزی پویا و نظریه بازیها را به اثبات رسانیده‌اند (Barquin et al., 2000).

پژوهش‌های متعددی نیز تأثیر زمان را در پیامدهای بازیکنان یا مدل‌های انحصاری دوجانبه مورد بررسی قرار داده‌اند. دری و همکاران رقابت بین دو زنجیره تأمین با درنظر گرفتن تقاضای وابسته به قیمت و زمان تدارک را تحت دو سناریوی رقابتی و یک سناریوی همکاری با استفاده از نظریه بازی‌ها مورد بررسی قرار داده‌اند. در مدل آن‌ها قیمت و زمان تدارک به عنوان متغیرهای تصمیم مورد تجزیه و تحلیل قرار گرفته‌اند (Dori et al., 2021). وو و همکاران، با رویکرد تئوری بازی‌ها رقابت دو زنجیره تأمین را با استراتژی‌های چانه زنی و استکلبرگ مورد واکاوی قرار داده‌اند. در مقاله آن‌ها مدل‌ها تک‌دوره‌ای و با دوره نامحدود فرض شده‌اند. از آنجاییکه در مدل ایشان بازه زمانی نامحدود است، از اثرگذاری انتقال موجودی از دوره‌ای به دوره دیگر چشمپوشی شده است (Wu et al., 2009). پگووان و همکاران یک مدل بازی برای تابع تقاضای وابسته به قیمت و زمان تدارک در رقابت بین دو شرکت را به عنوان یک تابع خطی توسعه داده‌اند. با این فرض که هر بنگاه دارای ظرفیت ثابت است و تقاضای هر بنگاه بطور خطی وابسته به قیمت و زمان تدارک تضمین شده است. عامل زمان یا تصمیم‌گیری درخصوص زمان تدارک در مدل تقاضای خطی آن‌ها برای محصول شرکت‌ها، جزء متغیرهای کنترلی در مسأله است (Pekgun et al., 2017). ما و همکاران یک مدل یکپارچه برای رقابت هزینه و زمان بین زنجیره‌های تأمین با مشتریان ناهمگون ارائه نموده‌اند. شرکت‌های موجود در مدل می‌توانند گزینه‌های زمانی مختلفی را برای تولید و خدمات خود به مشتریان حساس به زمان ارائه دهند. زمان عنصر اصلی در رقابت است و رقابت بین زنجیره تأمین را به بعد جدیدی یعنی زمان گسترش می‌دهد اما زمان در پژوهش آن‌ها نیز جزء متغیرهای تصمیم مسأله است (Ma et al., 2020).

در مورد وابستگی زمانی پیامدهای مربوط به بازیکنان در ماتریس پرداخت بازی نیز وارگا و همکاران، بازی‌های ماتریسی تحت محدودیت‌های زمانی را از طریق نظریه بازی‌های تکاملی مدل کرده‌اند. از دیدگاه آنها علاوه بر بازده، یک تعامل زوجی، پیامد دیگری برای شرکت کنندگان دارد. یعنی هر دو بازیکن باید برای شرکت در یک تعامل بعدی، مدتی صبر کنند تا آماده شوند. لذا با فرض این واقعیت با اختصاص یک سیستم دینامیکی پیوسته به هر بازی ماتریسی، توضیح می‌دهند که چگونه فراوانی فنوتایپ‌های مختلف در جمعیت، تکامل می‌یابد و درنهایت با مفهوم استراتژی‌های پایدار تکاملی^۴ (ESS) نقطه تعادل بازی را محاسبه می‌کنند (Varga et al., 2024). گارای و همکاران نیز یک بازی ماتریسی را تحت محدودیت‌های زمانی معرفی می‌کنند. در مفروضات آن‌ها هر تعامل زوجی دو نتیجه دارد: هر دو بازیکن یک بازده دریافت می‌کنند و نمی‌توانند بازی بعدی را

¹ Optimal control² Differential games³ Dynamic programming⁴ Evolutionary Stable Strategy

برای مدت زمان مشخصی انجام دهند. بنابراین مدل بازی با دو ماتریس تعریف می‌شود: یک ماتریس بازده و یک ماتریس مدت زمان متوسط. آنها در مدل خود تأثیر محدودیت‌های زمانی را با بازی معروف معمای زندانی بیان کرده و با تعمیم مفهوم ثبات تکاملی مینارد اسمیت به این دسته از بازی‌ها نشان می‌دهند که محدودیت‌های زمانی اضافی می‌توانند وجود استراتژی‌های پایدار تکاملی منحصر به فرد خالص و مختلط یا همزیستی دو ESS خالص را تضمین کنند (Garay et al., 2017).

کریوان و کرسمن در بازی‌های ماتریسی کلاسیک، تأثیری را که مدت زمان تعارض بر پرداخت‌ها می‌گذارند را با رویکرد جدیدی که هزینه‌های فرصت از دست رفته را در نظر می‌گیرد تحلیل نموده و تکامل همکاری و رقابت را از طریق دو مدل تأثیرگذار نظریه بازی‌های تکاملی، یعنی بازی شاهین و کبوتر و معمای زندانی در پویایی و تکامل جمعیت مورد بررسی قرار می‌دهند (Lrivan & Cressman, 2017). بروم و کریوان نیز با تعمیم تئوری محدودیت‌های زمانی بر روی مدل‌های بازی‌های تکاملی کریوان و کرسمن این واقعیت را در نظر می‌گیرند که رویدادهای مختلف، زمان‌های متفاوتی داشته و زمان‌های تعامل می‌تواند به صراحت به استراتژی‌های انتخاب شده بستگی داشته باشد. آنها نیز مفروضات خود را در مدلسازی و حل دو بازی معروف شاهین و کبوتر و معمای زندانی در نظر گرفته و نحوه تأثیر این فرضیات را در تعیین استراتژی‌های پایدار تکاملی مختلط بیان می‌کنند (Broom & Krivan, 2020).

در مورد مدل‌های انحصار دوجانبه پویای وابسته به زمان، آسکار و آلنویت در مقاله‌ای، پویایی پیچیده دو سیستم دوپولیستی کورنو را براساس پارامترهای وابسته به زمان با عقلانیت محدود و مکانیسم تلافی‌جویانه مورد مطالعه قرار داده‌اند. آنها تصمیم‌گیری در مورد میزان عرضه محصول را در دو مدل همکارانه و غیرهمکارانه تأخیری بررسی نموده و نتیجه می‌گیرند که پایداری سیستم منطقی با تأخیر، بهتر از آنهایی است که توسط استراتژی‌های تنبیهی اتخاذ می‌گردند (Askar & Alnowibet, 2016). بکهاف و ژانگ یک نسخه وابسته به زمان/اطلاعات از بازی‌های پویای کورنو را که تحت مفروضات خاصی ابداع شده است مدل کرده و وجود و منحصر به فرد بودن نقاط تعادل مدل فرضی را تحت عنوان تعادل پویای کورنو-نش اثبات می‌نمایند (Backhoff & Zhang, 2023). ایتزا و همکاران نیز در مقاله‌ای پایداری نقاط تعادل یک مدل دوقطبی کورنو را با فرار مالیاتی و تأخیر زمانی در یک چارچوب زمان پیوسته تحلیل نموده‌اند. در مدل ریاضی آنها هم نرخ هزینه نهایی و هم تأخیر زمانی به عنوان پارامترهای مسأله ایفای نقش می‌کنند. آنها نشان می‌دهند که اگر نرخ هزینه نهایی در بازه بسته معینی باشد، نقطه تعادل پایدار، مستقل از تأخیر است، در غیر اینصورت پایداری وابسته به تأخیر خواهد بود (Itza et al., 2020). در مورد وابستگی پیامدهای موجود در ماتریس پرداخت بازی به عاملی غیر از استراتژی انتخابی بازیکنان نیز، سام افلاکی بازی‌هایی را مورد بررسی قرار داده که در آنها پیامدهای موجود در ماتریس پرداخت بازی علاوه بر استراتژی انتخابی بازیکنان به یک متغیر تصمیم‌گیری نیز وابسته می‌باشد (Aflaki S, 2008).

همانطور که مشاهده شد در عموم پژوهش‌هایی که تاکنون پیرامون مسأله این پژوهش انجام شده است؛ عامل زمان به عنوان متغیرهای تصمیم مسأله و تحت کنترل بازیکنان بوده و با این فرض مدلسازی شده‌اند. در معدود پژوهش‌هایی نیز که زمان را به عنوان عاملی مستقل در پیامدهای بازیکنان در نظر گرفته‌اند در قالب بازی‌های دیفرانسیلی و یا از طریق نظریه بازی‌های تکاملی نسبت به مدلسازی و حل بازی اقدام نموده‌اند. اما تاکنون تأثیر عامل زمان به عنوان یک پارامتر مستقل در پیامدهای بازیکنان به این شیوه مورد تحلیل و بررسی قرار نگرفته است. پژوهش حاضر از تکنیک ترکیبی نظریه بازی و نظریه کنترل بهینه با رویکرد الگوریتمی برای مدل‌های انحصار دوقطبی استفاده نموده که به نوبه خود در روش مدلسازی و حل مدل شامل نوآوری است. در واقع در این پژوهش، با توجه به تغییر محیط و بازار با گذر زمان، از برنامه‌ریزی پویا و نظریه بازی برای تحلیل بازار و تعیین تصمیمات شرکت‌ها در بازار استفاده شده است. ساختار رقابت در این مطالعه شامل ویژگی‌های پویایی است که نمی‌توان آنها را به روش مطلوبی در مدل‌های کلاسیک و نرمال نمایش داد. در میان ساختارهای مختلف بازار، بررسی بازار انحصار دوجانبه یا دوقطبی، انتخاب شده است. در واقع هدف این پژوهش بدست آوردن مقادیر بهینه سطوح تولید و عرضه محصول انحصاری، در محیط رقابتی پویا در بازارهای دوقطبی است. در چنین محیط رقابتی، یک مدل کنترل بهینه توسعه داده شده است که از ابزارهای موجود در نظریه بازی‌ها یعنی تعادل نش و استکلبرگ جهت حداکثر کردن سود بازیکنان بهره می‌برد. برای تأیید اثربخشی تکنیک تلفیقی پیشنهادی، دو مثال عددی ارائه و روش‌شناسی برای حل آنها بکار گرفته شده است. مثال‌های عددی و تحلیل

نتایج، بینش مدیریتی مفید جهت انتخاب بهترین استراتژی‌ها برای بنگاه‌ها در محیط‌های مختلف بازار و عملیاتی را فراهم می‌آورد. مقایسه نتایج حاصل از ایده تعادل نش و استکلبرگ به تنهایی و روش تلفیقی پیشنهادی نیز، کارایی تکنیک ترکیبی را مشخص می‌نماید.

(ب) مدل کنترل بهینه گسسته

در این بخش کلاس ویژه‌ای از مدل‌های تصمیم‌بررسی و فرمول‌بندی شده که تابع هدف آنها به صورت الگوریتمی تعریف می‌شود. فرمول‌بندی ارائه شده برای مدل‌های تصمیم‌مزبور به همراه تکنیک‌های توسعه داده شده برای برنامه‌ریزی پویا امکان حل کلاس بزرگی از مسائل تصمیم را فراهم می‌آورد. مدل کنترل اصلی که برای چنین توسعه‌ای استفاده شده به شرح زیر است: سیستم پویای L با مجموعه‌ای از حالات $X \subseteq R^n$ را در هر لحظه از زمان $t=0,1,2,\dots$ در نظر بگیرید، حالت L برابر است با $x(t) \in X$ که $(x_1(t), x_2(t), \dots, x_n(t)) \in R^n$. دینامیک سیستم L به صورت زیر است:

$$x(t+1) = g_t(x(t), u(t)), \quad t = 0, 1, 2, \dots \quad \text{رابطه (۱)}$$

$$x(0) = x_s \quad \text{رابطه (۲)}$$

x_s ، نقطه شروع سیستم L است و $u(t) = (u_1(t), u_2(t), \dots, u_m(t)) \in R^m$ بردار متغیرهای تصمیم را نمایش می‌دهد. برای بردار پارامترهای کنترلی $(u(t), (t = 0, 1, 2, \dots))$ ، فضای تصمیم قابل قبول $U_t(x(t))$ اینگونه نمایش داده می‌شوند: $u(t) \in U_t(x(t)), t = 0, 1, 2, \dots$ رابطه (۳)

فرض می‌کنیم که در رابطه (۱)، توابع برداری

$$g_t(x(t), u(t)) = (g_t^1(x(t), u(t)), g_t^2(x(t), u(t)), \dots, g_t^n(x(t), u(t)))$$

به صورت یکتا توسط $x(t)$ و $u(t)$ در هر لحظه از زمان $t=0,1,2,\dots$ مشخص می‌شوند؛ بنابراین، $x(t+1)$ به صورت یکتا توسط $x(t)$ و $u(t)$ مشخص می‌شود. فرض کنید

$$x(0), x(1), \dots, x(t), \dots \quad \text{رابطه (۴)}$$

فرآیندی باشد با برداری مفروض از پارامترهای کنترلی $u(t), (t = 0, 1, 2, \dots)$ ، که بر طبق روابط (۱) - (۳) تولید شده است. برای هر حالت $(t = 0, 1, 2, \dots)$ ، $x(t)$ از فرایند (۴)، $f_t(x(t))$ به کمک فرمول بازگشتی زیر تعریف می‌شود

$$f_{t+1}(x(t+1)) = F_t(x(t), u(t), f_t(x(t))), \quad t = 0, 1, 2, \dots$$

$$f_0(x(0)) = f_0$$

f_0 ، نمایش مفروضی از حالت شروع $x(0)$ از سیستم L است و $F_t(\dots), t = 0, 1, 2, \dots$ توابع قراردادی هستند. در این مدل $f(x(t))$ عبارتست از هزینه‌های رفتن سیستم از حالت x_0 به $x(t)$.

مسئله بهینه‌سازی ۱: برای یک T مفروض، بردار متشکل از متغیرهای تصمیم $u(0), u(1), \dots, u(T-1)$ را به دست آورید که شرایط (۵) را برآورده و تابع هدف (۶) را کمینه سازد

$$\begin{cases} x(t+1) = g_t(x(t), u(t)) & t = 0, 1, 2, \dots, T-1; \\ x(0) = x_0 \quad x(T) = x_f & t = 0, 1, 2, \dots, T-1; \\ u(t) \in U_t(x(t)) & , \quad t = 0, 1, 2, \dots, T-1; \\ f_{t+1}(x(t+1)) = F_t(x(t), u(t), f_t(x(t))) \\ f_0(x(0)) = f_0 \end{cases} \quad \text{رابطه (۵)}$$

$$I_{x_0 x(T)}(u(t)) = f_T(x(T)) \quad \text{رابطه (۶)}$$

مسئله بهینه‌سازی ۲: برای T_1, T_2 مفروض، $T \in [T_1, T_2]$ و دنباله کنترلی $u(0), u(1), \dots, u(T-1)$ را مشخص کنید که شرایط (۵) را برآورده ساخته و تابع هدف (۶) را کمینه سازد. واضح است که حل بهینه مسئله ۲ را می‌توان با ثابت در نظر گرفتن پارامتر $T=T_1, T=T_1+1, \dots, T=T_2$ و تبدیل آن به مسئله ۱ به دست آورد. به کمک انتخاب مقدار بهینه حل‌های مسائلی از نوع ۱ با $T=T_1, T=T_1+1, \dots, T=T_2$ ، حل مربوط به مسئله ۲ را با $T \in [T_1, T_2]$ به دست می‌آوریم.

ج) الگوریتم حل

در این بخش یک روال کلی برای مشخص کردن جواب های بهینه از مسائل فرمول بندی شده با شرط زیر پیشنهاد می گردد:

$F_t(x, u, f), t = 0, 1, 2, \dots$ با توجه به آرگومان سوم f ، غیر نزولی هستند. بنابراین برای x و u ثابت، توابع

$F_t(x, u, f), t = 0, 1, 2, \dots$ شرط زیر را برآورده می سازند:

$$F_t(x, u, f') \leq F_t(x, u, f'') \quad \text{if } f' \leq f'' \quad \text{رابطه (۷)}$$

الگوریتم زیر، حل بهینه مسأله ۱ را تعیین می کند.

گام اول. قرار دهید

$$f_0^*(x(0)) = f_0; f_t^*(x(t)) = \infty; x(t) \in X_t; X_0 = \{x_0\}$$

گام دوم. برای $t = 1, 2, \dots, T$ مقادیر زیر را معین

$$X_{t+1} = \{x(t+1) \in X \mid x(t+1) = g_t(x(t), u(t)), x(t) \in X_t, u(t) \in U_t(x(t))\}$$

و برای هر $x(t+1) \in X_{t+1}$ تعیین کنید

$$f_{t+1}^*(x(t+1)) = \min \{F_t(x(t), u(t), f_t^*(x(t))) \mid x(t+1) = g_t(x(t), u(t)), x(t) \in X_t, u(t) \in U_t(x(t))\}$$

گام سوم. تعیین دنباله

$$x_T = x^*(T), x^*(T-1), x^*(T-2), \dots, x^*(1), x^*(0) = x_0$$

$$u^*(T-1), u^*(T-2), \dots, u^*(1), u^*(0)$$

که شرایط زیر را برآورده می سازند:

$$f_{T-\tau}^*(x^*(T-\tau)) = F_{T-\tau-1}(x^*(T-\tau-1), u^*(T-\tau-1), f_{T-\tau-1}^*(x(T-\tau-1)))$$

$$\tau = 0, 1, 2, \dots, T-1$$

در نهایت $u^*(T-1), u^*(T-2), \dots, u^*(1), u^*(0)$ حل بهینه مسأله ۱ را نمایش می دهد (Bellman & Kalaba, 1965); (Boltyanskii, 1973); (Krabs & Pickl, 2003).

قضیه: اگر $F_t(x, u, f), t = 0, 1, 2, \dots, T$ با توجه به آرگومان سوم f ، توابع غیر نزولی باشند؛ یعنی شرط (۷) را برآورده سازند، آنگاه الگوریتم فوق حل بهینه مسأله ۱ را تعیین می کند. بعلاوه، قسمت مقدم دلخواه $x^*(0), x^*(1), \dots, x^*(k)$ از مسیر بهینه $x^*(0), x^*(1), \dots, x^*(T)$ در مسأله ۱ نیز بهینه است (Lozovane & Pickl, 2006).

د) مدل کنترل بهینه بازی های ماتریسی

در این بخش رویکرد الگوریتمی ارائه شده برای حل مسائل کنترل بهینه گسسته، بر روی بازی های دو نفره با مجموع ثابت^۱ که در اصطلاح بازی های ماتریسی خوانده می شوند، پیاده سازی می گردد. حالتی را در نظر می گیریم که در آن پیامدهای هر بازیکن علاوه بر استراتژی انتخابی حریف به زمان یا دوره انجام بازی نیز وابسته می باشد که بایستی جهت ماکزیم نمودن سود در هر دوره زمانی از یک افق n پربودی، استراتژی مناسبی اتخاذ شود بگونه ای که سود مجموع این n دوره ماکزیم شود. از آنجاییکه بازی از دید بازیکن اول مورد بررسی قرار می گیرد لذا تابع هدف بازی بصورت زیر درمی آید:

$$\max_{u_t} \min_{v_t} \sum_{t=1}^n L_t(u_t, v_t)$$

$$\sum_{t=1}^n u_t \leq X, \quad \sum_{t=1}^n v_t \leq Y$$

$$u_t, v_t \geq 0 \quad \text{عدد صحیح}$$

بازی ماتریسی دو نفره زیر با مجموع ثابت C_t در مرحله t را در نظر بگیرید:

$$\begin{bmatrix} L_t(0,0) & \cdots & L_t(0,n_t) \\ \vdots & \ddots & \vdots \\ L_t(m_t,0) & \cdots & L_t(m_t,n_t) \end{bmatrix}$$

حال با توجه به ماتریس بازی فوق اقدام به مدلسازی برنامه ریزی پویای مسأله با حرکت رو به جلو می نماییم.

¹ Two-Person Constant-Sum Games

مرحله ۱. تعریف توابع و متغیرها

توابع برگشت بهین	متغیرهای حالت	متغیرهای تصمیم	دوره t
$f_t(X_t, Y_t)$	$X_t = \sum_{i=1}^t u_i$	u_t	بازیکن ۱
$C_t - f_t(X_t, Y_t)$	$Y_t = \sum_{i=1}^t v_i$	v_t	بازیکن ۲

مرحله ۲. تعیین حدود بالایی و پایینی متغیرهای حالت و تصمیم و گسسته کردن متغیرها بین حدود

$$\begin{aligned} 0 \leq u_t \leq X_t & \quad 0 \leq X_t \leq X \\ 0 \leq v_t \leq Y_t & \quad 0 \leq Y_t \leq Y \end{aligned}$$

مرحله ۳. تعیین معادله تکراری و شرط کمکی

$$\begin{aligned} f_1(X_1, Y_1) &= \max_{0 \leq u_1 \leq X_1} \min_{0 \leq v_1 \leq Y_1} L_1(u_1, v_1) \\ f_t(X_t, Y_t) &= \max_{0 \leq u_t \leq X_t} \min_{0 \leq v_t \leq Y_t} \{L_t(u_t, v_t) + f_{t-1}(X_t - u_t, Y_t - v_t)\} \end{aligned}$$

حال با طی گام‌های معرفی شده برای حل مسائل کنترل بهینه گسسته، بازی دونفره مجموع ثابت فوق، حل و استراتژی‌های بهینه هر مرحله تعیین می‌گردد. از آنجایی که رویکرد حرکت به جلو برای حل مسأله انتخاب شده است ابتدا مقادیر X و Y را برای متغیرهای حالت مرحله n ، فیکس می‌نماییم به عبارت دیگر $X_n = X$ و $Y_n = Y$ در نظر می‌گیریم. در گام بعد از شرط کمکی تعریف شده در فوق شروع نموده و با توجه به معادله تکراری مرحله به مرحله پیش می‌رویم و مقادیر بهینه تابع هدف در هر مرحله را می‌یابیم و در نهایت با رسیدن به مرحله n ، از انتها به ابتدا حرکت نموده، مسیر بهینه و دنباله تصمیم بهینه را تعیین می‌نماییم (Larson & Casti, 1978). جهت جامعیت مطلب رویکردهای ایستا و پویا در هر مرحله، مورد بررسی قرار می‌گیرد. در رویکرد ایستا، در هر مرحله یک بازی ایستا بین بازیکنان صورت می‌گیرد بدین معنا که بازیکنان در ابتدای هر مرحله استراتژی‌های خود را بصورت همزمان انتخاب می‌نمایند و هیچ بازیکنی از استراتژی انتخابی حریف آگاه نیست. نقطه تعادل به دست آمده در این حالت، تعادل نش است. اما در رویکرد پویا در هر مرحله ابتدا یک بازیکن به عنوان پیشرو، استراتژی خود را انتخاب می‌کند و بازیکن حریف آنرا مشاهده نموده و با توجه به آن، اقدام به انتخاب استراتژی خویش می‌نماید. تعادل به دست آمده در این حالت، تعادل استکلبرگ نامیده می‌شود. جهت آشنایی با روند حل چنین مسائلی از مثال عددی زیر استفاده شده- است.

ه) مثال عددی

فرض کنید که دو شرکت A و B محصولات مشابهی را بطور انحصاری به بازار عرضه می‌کنند. شرکت A به دنبال این است که چه میزان کالا را در هر دوره (با توجه به ظرفیت تولید n دوره ای) به بازار عرضه کند تا سود خود را در مجموع این n دوره ماکزیمم نماید. شرکت B وابسته به سندیکای^۱ D می‌باشد که مقید به درآمد خصوصی نیست و هدفش تنها براندازی شرکت A در این رقابت دوطرفه است لذا با استفاده از سرمایه D به راحتی زیان و سرمایه از دست رفته را جبران می‌کند. بعبارت دیگر هدف شرکت B این است که در این بازه n دوره ای شرکت A را از بازار محصول مورد نظر به در کرده و آنرا در انحصار خویش قرار دهد لذا برای بیرون کردن A از رقابت باید درآمد آنرا مینیمم نماید. فرض می‌گردد که بازی فوق در یک بازه یکساله صورت می‌گیرد و بازیکنان بایستی نسبت به میزان عرضه محصول در هر یک از فصول چهارگانه سال تصمیم‌گیری نمایند. از طرف دیگر فرض می‌شود که تقاضا برای محصول مورد نظر در هر دوره مقداری ثابت است در نتیجه مجموع سودی که از فروش محصول عاید دو شرکت می‌گردد برابر با مقداری ثابت خواهد بود لذا این مسأله در قالب یک بازی دونفره مجموع ثابت قابل مدل‌سازی می‌باشد. ظرفیت تولید سالیانه شرکت‌های A و B به ترتیب برابر با ۵ هزار و ۴ هزار واحد از محصول مورد نظر می‌باشد. تابع سود شرکت A در هر یک از فصول سال بصورت زیر برآورد گردیده :

$$\begin{aligned} L_1(u_1, v_1) &= 88u_1 - 12u_1^2 - 12u_1v_1 \\ L_2(u_2, v_2) &= 92u_2 - 12u_2^2 - 12u_2v_2 \\ L_3(u_3, v_3) &= 85u_3 - 12u_3^2 - 12u_3v_3 \end{aligned}$$

¹ Syndicate

$$L_4(u_4, v_4) = 90u_4 - 12u_4^2 - 12u_4v_4$$

که u_t و v_t در توابع فوق به ترتیب نشان دهنده میزان عرضه شرکت‌های A و B (برحسب هزار واحد) در دوره t می باشند. مجموع کالای عرضه شده به بازار در طی مراحل ۱ تا t توسط شرکت‌های A و B نیز به ترتیب با متغیرهای حالت X_t و Y_t تعریف می‌گردد. ماتریس پرداخت بازی طی مراحل ۱ تا ۴ با توجه به توابع سود تعریف شده در فوق در جدول (۱) آمده است.

جدول شماره (۱): ماتریس پرداخت بازی مراحل چهارگانه

ماتریس بازی مرحله ۱ با سود ثابت C_1						ماتریس بازی مرحله ۲ با سود ثابت C_2					
	۰	۱	۲	۳	۴		۰	۱	۲	۳	۴
۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
۱	۷۶	۶۴	۵۲	۴۰	۲۸	۱	۸۰	۶۸	۵۶	۴۴	۳۲
۲	۱۲۸	۱۰۴	۸۰	۵۶	۳۲	۲	۱۳۶	۱۱۲	۸۸	۶۴	۴۰
۳	۱۵۶	۱۲۰	۸۴	۴۸	۱۲	۳	۱۶۸	۱۳۲	۹۶	۶۰	۲۴
۴	۱۶۰	۱۱۲	۶۴	۱۶	-۳۲	۴	۱۷۶	۱۲۸	۸۰	۳۲	-۱۶
۵	۱۴۰	۸۰	۲۰	-۴۰	-۱۰۰	۵	۱۶۰	۱۰۰	۴۰	-۲۰	-۸۰
ماتریس بازی مرحله ۳ با سود ثابت C_3						ماتریس بازی مرحله ۴ با سود ثابت C_4					
	۰	۱	۲	۳	۴		۰	۱	۲	۳	۴
۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰
۱	۷۳	۶۱	۴۹	۳۷	۲۵	۱	۷۸	۶۶	۵۴	۴۲	۳۰
۲	۱۲۲	۹۸	۷۴	۵۰	۲۶	۲	۱۳۲	۱۰۸	۸۴	۶۰	۳۶
۳	۱۴۷	۱۱۱	۷۵	۳۹	۳	۳	۱۶۲	۱۲۶	۹۰	۵۴	۱۸
۴	۱۴۸	۱۰۰	۵۲	۴	-۴۴	۴	۱۶۸	۱۲۰	۷۲	۲۴	-۲۴
۵	۱۲۵	۶۵	۵	-۵۵	-۱۱۵	۵	۱۵۰	۹۰	۳۰	-۳۰	-۹۰

تعداد استکلبرگ

در این بخش فرض می‌کنیم که در هر مرحله یک بازی پویا انجام می‌گیرد یعنی ابتدا شرکت A کالای خود را به بازار عرضه نموده و شرکت B پس از مشاهده میزان عرضه A اقدام به عرضه کالا می‌نماید. حال با توجه به مفروضات فوق اقدام به حل مسأله می‌نماییم. از آنجاییکه بازی از دید بازیکن A مورد بررسی قرار می‌گیرد لذا استراتژی‌های ماکسیمین در هر مرحله بعنوان استراتژی‌های بهینه معرفی می‌گردند.

گام ۱. ابتدا متغیرهای حالت X_t و Y_t در مرحله n را با ظرفیت‌های تولید n دوره ای X و Y فیکس می‌گردد یعنی:

$$(X_4, Y_4) = (5, 4)$$

گام ۲. تعیین توابع برگشت بهین در هر مرحله

مرحله ۱. در ابتدا از شرط کمکی زیر شروع نموده و مقادیر بهینه تابع هدف در مرحله اول را می‌یابیم

$$f_1(X_1, Y_1) = \max_{0 \leq u_1 \leq 5} \{ \min_{0 \leq v_1 \leq 4} L_1(u_1, v_1) \}$$

مقادیر بهینه تابع هدف در مرحله یک به‌ازای مقادیر مختلف X_1 و Y_1 در جدول (۲) محاسبه گردیده است.

جدول شماره (۲): توابع برگشت بهین $f_1^*(X_1, Y_1)$

	۰	۱	۲	۳	۴
۰	۰	۰	۰	۰	۰
۱	(۰, ۰)	(۰, ۰)	(۰, ۰)	(۰, ۰)	(۰, ۰)
۲	(۱, ۰)	(۱, ۱)	(۱, ۲)	(۱, ۳)	(۱, ۴)
۳	(۲, ۰)	(۲, ۱)	(۲, ۲)	(۲, ۳)	(۲, ۴)
۴	(۳, ۰)	(۳, ۱)	(۳, ۲)	(۲, ۳)	(۲, ۴)

۴	۱۶۰	۱۲۰	۸۴	۵۶	۳۲
	(۴، ۰)	(۳، ۱)	(۳، ۲)	(۲، ۳)	(۲، ۴)
۵	۱۶۰	۱۲۰	۸۴	۵۶	۳۲
	(۴، ۰)	(۳، ۱)	(۳، ۲)	(۲، ۳)	(۲، ۴)

مرحله ۲. با توجه به معادله تکراری زیر مقادیر بهینه تابع هدف مرحله ۲ محاسبه می‌گردد:

$$f_2(X_2, Y_2) = \max_{0 \leq u_2 \leq X_2} \{ \min_{0 \leq v_2 \leq Y_2} (L_2(u_2, v_2) + f_1(X_2 - u_2, Y_2 - v_2)) \}$$

جدول شماره (۳): توابع برگشت بهین $f_2^*(X_2, Y_2)$

	۰	۱	۲	۳	۴
۰	۰ (۰، ۰)	۰ (۰، ۰)	۰ (۰، ۰)	۰ (۰، ۰)	۰ (۰، ۰)
۱	۸۰ (۱، ۰)	۶۸ (۱، ۱)	۵۶ (۱، ۲)	۴۴ (۱، ۳)	۳۲ (۱، ۴)
۲	۱۵۶ (۱، ۰)	۱۴۴ {(۱، ۲) $v_2=0,1$ }	۱۳۲ {(۱، ۲) $v_2=0,1,2$ }	۱۲۰ {(۱، ۲) $v_2=0,1,2,3$ }	۱۰۸ {(۱، ۲) $v_2=0,1,2,3,4$ }
۳	۲۱۲ (۲، ۰)	۱۸۸ (۲، ۱)	۱۶۴ (۲، ۲)	۱۴۰ (۲، ۳)	۱۱۶ (۲، ۴)
۴	۲۶۴ (۲، ۰)	۲۴۰ {(۲، ۲) $v_2=0,1$ }	۲۱۶ {(۲، ۲) $v_2=0,1,2$ }	۱۹۲ {(۲، ۲) $v_2=0,1,2,3$ }	۱۶۸ {(۲، ۲) $v_2=0,1,2,3,4$ }
۵	۲۹۶ (۳، ۰)	۲۶۰ (۳، ۱)	۲۲۴ (۳، ۲)	۱۹۲ (۲، ۰)	۱۶۸ {(۲، ۲) $v_2=0,1$ }

جهت آشنایی با نحوه محاسبه مقادیر موجود در جدول (۳)، یکی از مقادیر بطور دلخواه در زیر محاسبه گردیده است:

$$f_2(5,3) = \max_{0 \leq u_2 \leq 5} \{ \min_{0 \leq v_2 \leq 3} (L_2(u_2, v_2) + f_1(5 - u_2, 3 - v_2)) \}$$

$$f_2(5,3) = \max_{0 \leq u_2 \leq 5} \begin{cases} \min_{0 \leq v_2 \leq 3} (L_2(0, v_2) + f_1(5, 3 - v_2)) = \min(56, 84, 120, 160) = 56 \\ \min_{0 \leq v_2 \leq 3} (L_2(1, v_2) + f_1(4, 3 - v_2)) = \min(136, 152, 176, 204) = 136 \\ \min_{0 \leq v_2 \leq 3} (L_2(2, v_2) + f_1(3, 3 - v_2)) = \min(192, 196, 208, 220) = 192 \\ \min_{0 \leq v_2 \leq 3} (L_2(3, v_2) + f_1(2, 3 - v_2)) = \min(224, 212, 200, 188) = 188 \\ \min_{0 \leq v_2 \leq 3} (L_2(4, v_2) + f_1(1, 3 - v_2)) = \min(216, 180, 144, 108) = 108 \\ \min_{0 \leq v_2 \leq 3} (L_2(5, v_2) + f_1(0, 3 - v_2)) = \min(160, 100, 40, -20) = -20 \end{cases}$$

که با توجه به محاسبات فوق $f_2^*(5,3)=192$ می‌شود که استراتژی‌های بهینه $(u_2^*, v_2^*) = (2,0)$ را نتیجه می‌دهد.

مرحله ۳. مقادیر بهینه تابع هدف در مرحله ۳ با توجه به تابع تکراری زیر بازای ترکیبات مختلف X_3 و Y_3 محاسبه می‌گردد:

$$f_3(X_3, Y_3) = \max_{0 \leq u_3 \leq X_3} \{ \min_{0 \leq v_3 \leq Y_3} (L_3(u_3, v_3) + f_2(X_3 - u_3, Y_3 - v_3)) \}$$

مرحله ۴. از آنجاییکه در مرحله آخر $X_4=5$ و $Y_4=4$ می‌باشد لذا تنها کافیت $f_4(5,4)$ محاسبه شود.

$$f_4(5,4) = \max_{0 \leq u_4 \leq 5} \{ \min_{0 \leq v_4 \leq 4} (L_4(u_4, v_4) + f_1(5 - u_4, 4 - v_4)) \}$$

$$f_4(5,4) = \max_{0 \leq u_4 \leq 5} \begin{cases} \min_{0 \leq v_4 \leq 4} (L_4(0, v_4) + f_3(5, 4 - v_4)) = \min(241, 265, 289, 313, 337) = 241 \\ \min_{0 \leq v_4 \leq 4} (L_4(1, v_4) + f_3(4, 4 - v_4)) = \min(267, 279, 291, 303, 315) = 267 \\ \min_{0 \leq v_4 \leq 4} (L_4(2, v_4) + f_3(3, 4 - v_4)) = \min(313, 301, 289, 277, 265) = 265 \\ \min_{0 \leq v_4 \leq 4} (L_4(3, v_4) + f_3(2, 4 - v_4)) = \min(270, 246, 222, 198, 174) = 174 \\ \min_{0 \leq v_4 \leq 4} (L_4(4, v_4) + f_3(1, 4 - v_4)) = \min(200, 164, 128, 92, 56) = 56 \\ \min_{0 \leq v_4 \leq 4} (L_4(5, v_4) + f_3(0, 4 - v_4)) = \min(150, 90, 30, -30, -90) = -90 \end{cases}$$

محاسبات فوق $f_4^*(5,4)=267$ و استراتژی‌های بهینه $(u_4^*, v_4^*)=(1,0)$ را نتیجه می‌دهند.

گام ۳. تعیین دنباله تصمیم بهینه و مسیر بهینه

حال که تابع برگشت بهین در مرحله ۴ یا عبارت دیگر سود حاصله در مجموع ۴ دوره بدست آمد، از انتها به ابتدا حرکت کرده دنباله تصمیم بهینه و مسیر بهینه را تعیین می‌نماییم.

جدول شماره (۴): دنباله تصمیم بهینه و مسیر بهینه

مرحله	مسیر بهینه	دنباله تصمیم بهینه
۴	$(X_4, Y_4) = (5, 4)$	$(u_4^*, v_4^*) = (1, 0)$
۳	$(X_3, Y_3) = (X_4 - u_4^*, Y_4 - v_4^*) = (4, 4)$	$(u_3^*, v_3^*) = (1, 0)$
۲	$(X_2, Y_2) = (X_3 - u_3^*, Y_3 - v_3^*) = (3, 4)$	$(u_2^*, v_2^*) = (2, 4)$
۱	$(X_1, Y_1) = (X_2 - u_2^*, Y_2 - v_2^*) = (1, 0)$	$(u_1^*, v_1^*) = (1, 0)$

از آنجا که شرکت A در ابتدای هر دوره قبل از شرکت B جهت میزان عرضه کالا در بازار تصمیم‌گیری می‌نماید لذا کل ظرفیت تولید ۴ دوره‌ای را در دوره ۳ به شرکت B تحویل می‌نماید. جهت نشان دادن کارایی تکنیک معرفی شده، دنباله تصمیم بهینه و مسیر بهینه‌ای را که تعادل استکلبرگ به تنهایی و بدون کاربرد تکنیک برنامه ریزی پویا بدست می‌دهد، محاسبه و با نتایج فوق مقایسه می‌گردد.

جدول شماره (۵): مسیر بهینه و دنباله تصمیم بهینه حاصل از تعادل نش با حرکت رو به جلو

مرحله	مسیر بهینه	دنباله تصمیم بهینه	تابع برگشت بهین
۱	$(X_1, Y_1) = (5, 4)$	$(u_1^*, v_1^*) = (2, 4)$	$f_1^*(5, 4) = 32$
۲	$(X_2, Y_2) = (3, 0)$	$(u_2^*, v_2^*) = (3, 0)$	$f_2^*(3, 0) = 32 + 168 = 200$
۳	$(X_3, Y_3) = (0, 0)$	$(u_3^*, v_3^*) = (0, 0)$	$f_3^*(0, 0) = 200 + 0 = 200$
۴	$(X_4, Y_4) = (0, 0)$	$(u_4^*, v_4^*) = (0, 0)$	$f_4^*(0, 0) = 200 + 0 = 200$

جدول شماره (۶): مسیر بهینه و دنباله تصمیم بهینه حاصل از تعادل نش با حرکت رو به عقب

مرحله	مسیر بهینه	دنباله تصمیم بهینه	تابع برگشت بهین
۴	$(X_4, Y_4) = (5, 4)$	$(u_4^*, v_4^*) = (2, 4)$	$f_4^*(5, 4) = 36$
۳	$(X_3, Y_3) = (3, 0)$	$(u_3^*, v_3^*) = (3, 0)$	$f_3^*(3, 0) = 36 + 147 = 183$
۲	$(X_2, Y_2) = (0, 0)$	$(u_2^*, v_2^*) = (0, 0)$	$f_2^*(0, 0) = 183 + 0 = 183$
۱	$(X_1, Y_1) = (0, 0)$	$(u_1^*, v_1^*) = (0, 0)$	$f_1^*(0, 0) = 183 + 0 = 183$

همانطور که مشاهده می‌گردد سود بدست آمده از دنباله تصمیم بهینه و مسیر بهینه تکنیک ترکیبی، برابر ۲۶۷ و سود حاصله از رویکرد نظریه بازی صرف، حداکثر برابر ۲۰۰ است.

تعادل نش

در این قسمت فرض می‌شود که در هر مرحله یک بازی ایستا بین بازیکنان صورت می‌گیرد یعنی در هر مرحله بازیکنان بطور همزمان استراتژی خویش را انتخاب می‌کنند و هیچ بازیکنی از استراتژی انتخابی حریف مطلع نمی‌باشد. همانطور که مشاهده گردید در رویکرد پویا استراتژی ماکسیمین در هر مرحله به عنوان استراتژی بهینه انتخاب می‌شد که همواره به استراتژی محض منجر می‌گردید. در رویکرد ایستا استراتژی‌های بهینه هر مرحله از طریق محاسبه تعادل نش ماتریس پرداخت بازی حاصل از معادلات تکراری همان مرحله محاسبه می‌گردد لذا این امکان وجود دارد که در بعضی مراحل ماتریس پرداخت بازی حاصل از معادلات تکراری مرحله مزبور فاقد نقطه تعادل محض باشد که در اینصورت بایستی تعادل نش مختلط بازی را محاسبه نمود. بطور مثال جهت محاسبه مقدار بهینه تابع هدف در مرحله t یا عبارت دیگر محاسبه تابع برگشت بهین $f_t^*(X_t, Y_t)$ ، بایستی نقطه تعادل نش ماتریس پرداخت زیر را محاسبه نمود:

$$\begin{bmatrix} L_t(0,0) + f_{t-1}(X_t - 0, Y_t - 0) & \cdots & L_t(0, Y_t) + f_{t-1}(X_t, 0) \\ \vdots & \ddots & \vdots \\ L_t(X_t, 0) + f_{t-1}(0, Y_t) & \cdots & L_t(X_t, Y_t) + f_{t-1}(0, 0) \end{bmatrix}$$

اگر ماتریس پرداخت فوق دارای نقطه تعادل نش محض باشد که محاسبات و نحوه پیشروی مراحل دقیقاً با رویکرد پویا مطابقت

خواهد داشت اما اگر ماتریس پرداخت فوق فاقد نقطه تعادل نش محض باشد باید نقطه تعادل نش مختلط را محاسبه نمود. برای حل ماتریس پرداخت فوق و یافتن نقطه تعادل نش، می‌توان از روش برنامه‌ریزی خطی استفاده نمود (Hamdy Taha, 1976). حال فرض کنید که حل بهینه ماتریس پرداخت فوق به روش LP به صورت زیر باشد:

$$(P^*, Q^*, V^*) = ((p_1^t, p_2^t, \dots, p_{X_t}^t), (q_1^t, q_2^t, \dots, q_{Y_t}^t), f_t(X_t, Y_t))$$

لذا نتیجه می‌شود که

$$(u_t^*, v_t^*) = \left(\sum_{i=0}^{X_t} i p_i^t, \sum_{j=0}^{Y_t} j q_j^t \right)$$

(و) مدل کنترل بهینه بازیهای دوماتریسی

رویکرد الگوریتمی ارائه شده تا به اینجا بر روی بازیهای دونفره با مجموع ثابت پیاده سازی گردید اما در مورد بازیهای دونفره با مجموع غیر ثابت^۱ یا بازیهای دوماتریسی از آنجا که در هر یک از درایه های ماتریس پرداخت شاهد یک زوج پیامد هستیم لذا به دو شرط کمکی و دو معادله تکراری در حل مسأله نیاز است. از آنجاییکه تابع هدف هر بازیکن علاوه بر متغیر تصمیم خویش به متغیر تصمیم بازیکن دیگر یا عبارت دیگر به استراتژی اتخاذی حریف نیز وابسته می‌باشد، امکان حل مجزای دو تابع هدف وجود ندارد که بایستی بهینه سازی توأم توابع هدف و نحوه پیشروی مراحل مورد بررسی قرارگیرد. در این بخش رویکرد الگوریتمی ارائه شده برای حل مسائل کنترل بهینه گسسته، بر روی بازی های دو ماتریسی پیاده سازی می‌گردد. برخلاف بازیهای ماتریسی که یک بازیکن در پی ماکزیمم نمودن سود و بازیکن دیگر در پی مینیمم نمودن زیان خویش است، در بازیهای دوماتریسی علی رغم تأثیر متقابل تصمیمات طرفین در پیامد دریافتی، هر دو بازیکن در پی ماکزیمم نمودن سود دریافتی از انجام بازی هستند لذا تابع هدف بازی بصورت زیر در می‌آید:

$$\begin{aligned} & (\max_{u_t} \sum_{t=1}^n L_t(u_t, v_t), \max_{v_t} \sum_{t=1}^n K_t(u_t, v_t)) \\ & \sum_{t=1}^n u_t \leq X, \quad \sum_{t=1}^n v_t \leq Y \\ & u_t, v_t \geq 0 \quad \text{عدد صحیح} \end{aligned}$$

بازی دوماتریسی زیر در مرحله t را در نظر بگیرید:

$$\begin{bmatrix} L_t(0,0), K_t(0,0) & \cdots & L_t(0, n_t), K_t(0, n_t) \\ \vdots & \ddots & \vdots \\ L_t(m_t, 0), K_t(m_t, 0) & \cdots & L_t(m_t, n_t), K_t(m_t, n_t) \end{bmatrix}$$

حال با توجه به ماتریس بازی فوق اقدام به مدلسازی برنامه ریزی پویای مسأله با حرکت رو به جلو می‌گردد.

مرحله ۱. تعریف توابع و متغیرها

توابع برگشت بهین	متغیرهای حالت	متغیرهای تصمیم	دوره t
$f_t(X_t, Y_t)$	$X_t = \sum_{i=1}^t u_i$	u_t	بازیکن ۱
$g_t(X_t, Y_t)$	$Y_t = \sum_{i=1}^t v_i$	v_t	بازیکن ۲

مرحله ۲. تعیین حدود بالایی و پایینی متغیرهای حالت و تصمیم و گسسته کردن متغیرها بین حدود

$$\begin{aligned} 0 \leq X_t \leq X & \quad 0 \leq u_t \leq X_t \\ 0 \leq Y_t \leq Y & \quad 0 \leq v_t \leq Y_t \end{aligned}$$

مرحله ۳. تعیین معادله تکراری و شروط کمکی

$$\begin{aligned} f_1(X_1, Y_1) &= \max_{0 \leq u_1 \leq X_1} L_1(u_1, v_1) \\ g_1(X_1, Y_1) &= \max_{0 \leq v_1 \leq Y_1} K_1(u_1, v_1) \end{aligned}$$

¹ Two-Person Nonconstant-Sum Games

$$f_t(X_t, Y_t) = \max_{0 \leq u_t \leq X_t} \{L_t(u_t, v_t) + f_{t-1}(X_t - u_t, Y_t - v_t)\}$$

$$g_t(X_t, Y_t) = \max_{0 \leq v_t \leq Y_t} \{K_t(u_t, v_t) + g_{t-1}(X_t - u_t, Y_t - v_t)\}$$

تنها مشکل موجود در حل مدل ارائه شده، بهینه سازی همزمان دو تابع هدف فوق است. از آنجاییکه تابع هدف هر بازیکن علاوه بر متغیر تصمیم خویش به متغیر تصمیم بازیکن دیگر یا عبارت دیگر به استراتژی اتخاذی حریف نیز وابسته می باشد، امکان حل مجزای دو تابع هدف فوق وجود ندارد لذا جهت بهینه سازی توأم توابع هدف از مفهوم تعادل نش استفاده می گردد به اینصورت که از شروط کمکی شروع نموده و با معادلات تکراری با متد برنامه ریزی پویا پیش می رویم و بهینه توأم توابع هدف در هر یک از مراحل را با یافتن تعادل نش ماتریس بازی حاصل از توابع تکراری می یابیم. بطور مثال نقطه تعادل نش بازی دوماتریسی زیر، مقدار بهینه توابع هدف در مرحله t یا عبارت دیگر $f_t^*(X_t, Y_t)$ و $g_t^*(X_t, Y_t)$ را بدست می دهد.

$$\begin{bmatrix} L_t(0,0) + f_{t-1}(X_t - 0, Y_t - 0) & \cdots & L_t(0, Y_t) + f_{t-1}(X_t, 0) \\ \vdots & \ddots & \vdots \\ L_t(X_t, 0) + f_{t-1}(0, Y_t) & \cdots & L_t(X_t, Y_t) + f_{t-1}(0, 0) \\ K_t(0,0) + g_{t-1}(X_t - 0, Y_t - 0) & \cdots & K_t(0, Y_t) + g_{t-1}(X_t, 0) \\ \vdots & \ddots & \vdots \\ K_t(X_t, 0) + g_{t-1}(0, Y_t) & \cdots & K_t(X_t, Y_t) + g_{t-1}(0, 0) \end{bmatrix}$$

درنهایت و با رسیدن به مرحله n ، از انتها به ابتدا حرکت نموده و دنباله تصمیم بهینه و مسیر بهینه را می یابیم.

(ز) مثال عددی

مدل انحصار دو قطبی کورنو^۱ را در نظر بگیرید که در آن دو شرکت A و B بازار مربوط به یک محصول خاص را کنترل می کنند (Straffin, 1993). هر دو شرکت این هدف را دنبال می کنند که چه میزان کالا را در هر دوره با توجه به ظرفیت تولید n دوره ای به بازار عرضه نموده تا سود خویش را در مجموع این n دوره حداکثر نمایند. مجدداً فرض می شود که سیکل زندگی محصول مورد نظر یکساله بوده و شرکتها بایستی نسبت به میزان تولید محصول در هر یک از فصول چهارگانه سال تصمیم گیری نمایند. علاوه بر این فرض می گردد که ظرفیت تولید سالیانه شرکت اول برابر با ۷ هزار واحد و شرکت دوم برابر با ۵ هزار واحد از محصول مورد نظر است. متغیرهای مسأله را بصورت زیر تعریف می کنیم:

u_t : میزان عرضه شرکت A (برحسب هزار) در دوره t P_t : قیمت هر قلم محصول در دوره t
 v_t : میزان عرضه شرکت B (برحسب هزار) در دوره t $L_t(u_t, v_t)$: سود (برحسب هزار دلار) شرکت A در دوره t
 C_{t1} : هزینه متوسط تولید هر قلم کالا برای شرکت A در دوره t $K_t(u_t, v_t)$: سود (برحسب هزار دلار) شرکت B در دوره t
 C_{t2} : هزینه متوسط تولید هر قلم کالا برای شرکت B در دوره t
توابع سود شرکتها در هر یک از فصول سال در جدول (۷) برآورد گردیده که با توجه به توابع سود تعریف شده، ماتریس پرداخت بازی مراحل چهارگانه را محاسبه و نتیجه در جدول (۸) آمده است.

جدول شماره (۷): توابع هزینه، قیمت و سود شرکتها A و B

مرحله	توابع هزینه	تابع قیمت	توابع سود
۱	$C_{11} = 64 - 4u_1 + u_1^2$ $C_{12} = 80 - 4v_1 + v_1^2$	$P_1 = 160 - 8(u_1 + v_1)$	$L_1(u_1, v_1) = (P_1 - C_{11})u_1 = 96u_1 - 4u_1^2 - u_1^3 - 8u_1v_1$ $K_1(u_1, v_1) = (P_1 - C_{12})v_1 = 80v_1 - 4v_1^2 - v_1^3 - 8u_1v_1$
۲	$C_{21} = 65 - 4u_2 + u_2^2$ $C_{22} = 82 - 4v_2 + v_2^2$	$P_2 = 165 - 8(u_2 + v_2)$	$L_2(u_2, v_2) = (P_2 - C_{21})u_2 = 100u_2 - 4u_2^2 - u_2^3 - 8u_2v_2$ $K_2(u_2, v_2) = (P_2 - C_{22})v_2 = 83v_2 - 4v_2^2 - v_2^3 - 8u_2v_2$
۳	$C_{31} = 66 - 4u_3 + u_3^2$ $C_{32} = 82 - 4v_3 + v_3^2$	$P_3 = 158 - 8(u_3 + v_3)$	$L_3(u_3, v_3) = (P_3 - C_{31})u_3 = 92u_3 - 4u_3^2 - u_3^3 - 8u_3v_3$ $K_3(u_3, v_3) = (P_3 - C_{32})v_3 = 76v_3 - 4v_3^2 - v_3^3 - 8u_3v_3$
۴	$C_{41} = 67 - 4u_4 + u_4^2$ $C_{42} = 85 - 4v_4 + v_4^2$	$P_4 = 157 - 8(u_4 + v_4)$	$L_4(u_4, v_4) = (P_4 - C_{41})u_4 = 90u_4 - 4u_4^2 - u_4^3 - 8u_4v_4$ $K_4(u_4, v_4) = (P_4 - C_{42})v_4 = 72v_4 - 4v_4^2 - v_4^3 - 8u_4v_4$
جدول شماره (۸): ماتریس پرداخت بازی مراحل چهارگانه			
ماتریس بازی مرحله ۱		ماتریس بازی مرحله ۲	

¹ Cournot's Duopoly model

۴	۳	۲	۱	۰	۴	۳	۲	۱	۰
۰,۲۰۴	۰,۱۸۶	۰,۱۴۲	۰,۷۸	۰,۰	۰,۱۹۲	۰,۱۷۷	۰,۱۳۶	۰,۷۵	۰,۰
۶۳,۱۷۲	۷۱,۱۶۲	۷۹,۱۲۶	۸۷,۷۰	۹۵,۰	۵۹,۱۶۰	۶۷,۱۵۳	۷۵,۱۲۰	۸۳,۶۷	۹۱,۰
۱۱۲,۱۴۰	۱۲۸,۱۳۸	۱۴۴,۱۱۰	۱۶۰,۶۲	۱۷۶,۰	۱۰۴,۱۲۸	۱۲۰,۱۲۹	۱۳۶,۱۰۴	۱۵۲,۵۹	۱۶۸,۰
۱۴۱,۱۰۸	۱۶۵,۱۱۴	۱۸۹,۹۴	۲۱۳,۵۴	۲۳۷,۰	۱۲۹,۹۶	۱۵۳,۱۰۵	۱۷۷,۸۸	۲۰۱,۵۱	۲۲۵,۰
۱۴۴,۷۶	۱۷۶,۹۰	۲۰۸,۷۸	۲۴۰,۴۶	۲۷۲,۰	۱۲۸,۶۴	۱۶۰,۸۱	۱۹۲,۷۲	۲۲۴,۴۳	۲۵۶,۰
۱۱۵,۴۴	۱۵۵,۶۶	۱۹۵,۶۲	۲۳۵,۳۸	۲۷۵,۰	۹۵,۳۲	۱۳۵,۵۷	۱۷۵,۵۶	۲۱۵,۳۵	۲۵۵,۰
ماتریس بازی مرحله ۴					ماتریس بازی مرحله ۳				
۴	۳	۲	۱	۰	۴	۳	۲	۱	۰
۰,۱۶۰	۰,۱۵۳	۰,۱۲۰	۰,۶۷	۰,۰	۰,۱۷۶	۰,۱۶۵	۰,۱۲۸	۰,۷۱	۰,۰
۵۳,۱۲۸	۶۱,۱۲۹	۶۹,۱۰۴	۷۷,۵۹	۸۵,۰	۵۵,۱۴۴	۶۳,۱۴۱	۷۱,۱۱۲	۷۹,۶۳	۸۷,۰
۹۲,۹۶	۱۰۸,۱۰۵	۱۲۴,۸۸	۱۴۰,۵۱	۱۵۶,۰	۹۶,۱۱۲	۱۱۲,۱۱۷	۱۲۸,۹۶	۱۴۴,۵۵	۱۶۰,۰
۱۱۱,۶۴	۱۳۵,۸۱	۱۵۹,۷۲	۱۸۳,۴۳	۲۰۷,۰	۱۱۷,۸۰	۱۴۱,۹۳	۱۶۵,۸۰	۱۸۹,۴۷	۲۱۳,۰
۱۰۴,۳۲	۱۳۶,۵۷	۱۶۸,۵۶	۲۰۰,۳۵	۲۳۲,۰	۱۱۲,۴۸	۱۴۴,۶۹	۱۷۶,۶۴	۲۰۸,۳۹	۲۴۰,۰
۶۵,۰	۱۰۵,۳۳	۱۴۵,۴۰	۱۸۵,۲۷	۲۲۵,۰	۷۵,۱۶	۱۱۵,۴۵	۱۵۵,۴۸	۱۹۵,۳۱	۲۳۵,۰

حل بازی با تعادل نش محض

گام ۱. ابتدا متغیرهای حالت X_t و Y_t در مرحله n با ظرفیت‌های تولید n دوره ای X و Y فیکس می‌شود یعنی:

$$(X_4, Y_4) = (7, 5)$$

گام ۲. تعیین توابع برگشت بهین در هر مرحله

با استفاده از شروط کمکی و معادلات تکراری زیر مقادیر بهینه تابع هدف بازیکنان در هر مرحله محاسبه می‌گردد:

$$f_1(X_1, Y_1) = \max_{0 \leq u_1 \leq X_1} L_1(u_1, v_1)$$

$$g_1(X_1, Y_1) = \max_{0 \leq v_1 \leq Y_1} K_1(u_1, v_1)$$

$$f_t(X_t, Y_t) = \max_{0 \leq u_t \leq X_t} \{L_t(u_t, v_t) + f_{t-1}(X_t - u_t, Y_t - v_t)\}$$

$$g_t(X_t, Y_t) = \max_{0 \leq v_t \leq Y_t} \{K_t(u_t, v_t) + g_{t-1}(X_t - u_t, Y_t - v_t)\}$$

جهت آشنایی با نحوه محاسبه مقادیر بهینه توابع هدف در هر مرحله، درایه (۵,۴) از ماتریس توابع برگشت بهین $f_2^*(X_2, Y_2)$ بطور دلخواه در زیر محاسبه شده است.

۴	۳	۲	۱	۰
۲۵۶,۲۰۴	۲۲۴,۲۲۹	۱۹۲,۲۱۴	۱۶۰,۱۵۹	۱۶۰,۸۱
۳۱۹,۱۷۲	۲۹۵,۲۰۵	۲۷۱,۱۹۸	۲۴۷,۱۵۱	۲۵۵,۸۱
۳۳۷,۱۴۰	۳۳۹,۱۸۹	۳۲۱,۱۹۸	۳۱۳,۱۶۷	۳۲۹,۱۰۵
۳۰۹,۱۰۸	۳۱۷,۱۷۳	۳۲۵*,۱۹۸*	۳۳۳,۱۸۳	۳۵۷,۱۲۹
۲۳۵,۷۶	۲۵۹,۱۵۷	۲۸۳,۱۹۸	۳۰۷,۱۹۹	۳۳۱,۱۶۰
۱۱۵,۴۴	۱۵۵,۱۴۱	۱۹۵,۱۹۸	۲۳۵,۲۱۵	۲۷۵,۱۹۲

نقطه تعادل نش محاسبه شده در ماتریس پرداخت فوق، مقادیر بهینه توابع هدف و استراتژی‌های بهینه زیر را بدست می‌دهد:

$$(f_2^*(5,4), g_2^*(5,4)) = (325, 198) \rightarrow (u_2^*, v_2^*) = (3, 2)$$

از آنجاییکه در مرحله چهارم $X_4=7$ و $Y_4=5$ می‌باشد لذا در مرحله آخر تنها محاسبه $f_4(7,5)$ و $g_4(7,5)$ کفایت می‌کند.

ماتریس پرداخت حاصل از توابع تکراری $f_4(7,5)$ و $g_4(7,5)$ ، نقاط تعادل نش و استراتژی‌های بهینه زیر را بدست می‌دهد:

$$(f_4^*(7,5), g_4^*(7,5)) = (517, 283) \rightarrow (u_4^*, v_4^*) = (1, 1)$$

$$(f_4^*(7,5), g_4^*(7,5)) = (519, 288) \rightarrow (u_4^*, v_4^*) = (2, 1)$$

گام ۳. تعیین دنباله تصمیم بهینه و مسیر بهینه

حال که مقادیر بهینه توابع هدف در هر یک از مراحل محاسبه گردید، از انتها به ابتدا حرکت نموده دنباله تصمیم بهینه و

مسیر بهینه را مشخص می‌نماییم.

جدول شماره (۹): دنباله تصمیمات و مسیرهای بهینه

مسیر بهینه ۱			
مرحله	مسیر بهینه	دنباله تصمیم بهینه	توابع بر گشت بهین
۴	$(X_4, Y_4) = (7, 5)$	$(u_4^*, v_4^*) = (1, 1)$	$f_4(7, 5) = 517$
۳	$(X_3, Y_3) = (6, 4)$	$(u_3^*, v_3^*) = (2, 1)$	
۲	$(X_2, Y_2) = (4, 3)$	$(u_2^*, v_2^*) = (2, 2)$	$g_4(7, 5) = 283$
۱	$(X_1, Y_1) = (2, 1)$	$(u_1^*, v_1^*) = (2, 1)$	
مسیر بهینه ۲			
مرحله	مسیر بهینه	دنباله تصمیم بهینه	توابع بر گشت بهین
۴	$(X_4, Y_4) = (7, 5)$	$(u_4^*, v_4^*) = (2, 1)$	$f_4(7, 5) = 519$
۳	$(X_3, Y_3) = (5, 4)$	$(u_3^*, v_3^*) = (2, 1)$	
۲	$(X_2, Y_2) = (3, 3)$	$(u_2^*, v_2^*) = (2, 1)$	$g_4(7, 5) = 288$
۱	$(X_1, Y_1) = (1, 2)$	$(u_1^*, v_1^*) = (1, 2)$	

جهت نشان دادن کارایی تکنیک معرفی شده، دنباله تصمیم بهینه و مسیر بهینه‌ای را که تعادل نش به تنهایی و بدون کاربرد تکنیک برنامه‌ریزی پویا بدست می‌دهد، محاسبه و با نتایج فوق مقایسه می‌گردد.

جدول شماره (۱۰): مسیر بهینه و دنباله تصمیم بهینه حاصل از تعادل نش با حرکت رو به جلو

مرحله	مسیر بهینه	دنباله تصمیم بهینه	تابع برگشت بهین
۱	$(X_1, Y_1) = (7, 5)$	$(u_1^*, v_1^*) = (4, 3)$	$(f_1(7, 5), g_1(7, 5)) = (160, 81)$
۲	$(X_2, Y_2) = (3, 2)$	$(u_2^*, v_2^*) = (3, 2)$	$(f_2(3, 2), g_2(3, 2)) = (349, 175)$
۳	$(X_3, Y_3) = (0, 0)$	$(u_3^*, v_3^*) = (0, 0)$	$(f_3(0, 0), g_3(0, 0)) = (349, 175)$
۴	$(X_4, Y_4) = (0, 0)$	$(u_4^*, v_4^*) = (0, 0)$	$(f_4(0, 0), g_4(0, 0)) = (349, 175)$

جدول شماره (۱۱): مسیر بهینه و دنباله تصمیم بهینه حاصل از تعادل نش با حرکت رو به عقب

مرحله	مسیر بهینه	دنباله تصمیم بهینه	تابع برگشت بهین
۴	$(X_4, Y_4) = (7, 5)$	$(u_4^*, v_4^*) = (4, 3)$	$(f_4(7, 5), g_4(7, 5)) = (136, 57)$
۳	$(X_3, Y_3) = (3, 2)$	$(u_3^*, v_3^*) = (3, 2)$	$(f_3(3, 2), g_3(3, 2)) = (301, 137)$
۲	$(X_2, Y_2) = (0, 0)$	$(u_2^*, v_2^*) = (0, 0)$	$(f_2(0, 0), g_2(0, 0)) = (301, 137)$
۱	$(X_1, Y_1) = (0, 0)$	$(u_1^*, v_1^*) = (0, 0)$	$(f_1(0, 0), g_1(0, 0)) = (301, 137)$

مشاهده می‌گردد که نتایج بدست آمده از تکنیک ترکیبی، نسبت به استفاده صرف از تعادل نش، نتایج بسیار مطلوبتری را بدست می‌دهد.

حل بازی با تعادل نش مختلط

همانطور که ملاحظه گردید در مثال فوق ماتریس پرداخت بازی حاصل از معادلات تکراری در تمامی مراحل دارای نقطه تعادل نش محض بود اما اگر ماتریس پرداخت بازی حاصل از معادلات تکراری در برخی مراحل فاقد نقطه تعادل نش محض باشد محاسبات و روند پیشروی مراحل تغییر چندانی نمی‌کند و تنها بایستی نقطه تعادل نش مختلط را محاسبه نمود. فرض کنید که ماتریس پرداخت بازی حاصل از معادلات تکراری مرحله t فاقد نقطه تعادل نش محض باشد. جهت سهولت نمایش معادلات مربوطه که به محاسبه تعادل نش مختلط منجر می‌گردد زوج پیامدهای موجود در ماتریس پرداخت با (a_{ij}^t, b_{ij}^t) نمایش داده می‌شود که در آن :

$$a_{ij}^t = L_t(i, j) + f_{t-1}(X_t - i, Y_t - j)$$

$$b_{ij}^t = K_t(i, j) + g_{t-1}(X_t - i, Y_t - j)$$

روش محاسبه پیامدها و استراتژی‌های مختلط بازیکنان بدین صورت است که با تخصیص احتمالات $P^t = (p_1^t, p_2^t, \dots, p_{X_t}^t)$ به سطرها و $Q^t = (q_1^t, q_2^t, \dots, q_{Y_t}^t)$ به ستونها، دریافتی مورد انتظار بازیکن A بصورت

$$\pi_A(P^t, Q^t) = \sum_{i=0}^{X_t} \sum_{j=0}^{Y_t} p_i^t a_{ij}^t q_j^t$$

و دریافتی موردانتظار بازیکن B به صورت زیر در می آید

$$\pi_B(P^t, Q^t) = \sum_{i=0}^{X_t} \sum_{j=0}^{Y_t} p_i^t b_{ij}^t q_j^t$$

تعالیل نش مختلط بازی می تواند از حل معادلات مشتق جزئی زیر حاصل گردد (Kelly, 2003).

$$\begin{cases} \frac{\partial \pi_A}{\partial p_i^t} = 0 & \text{for } i = 0, 1, 2, \dots, X_t \\ \frac{\partial \pi_B}{\partial q_j^t} = 0 & \text{for } j = 0, 1, 2, \dots, Y_t \end{cases}$$

که در معادلات فوق خواهیم داشت :

$$\sum_{i=0}^{X_t} p_i^t = 1, 0 \leq p_i^t \leq 1 \quad \forall i = 0, 1, 2, \dots, X_t$$

$$\sum_{j=0}^{Y_t} q_j^t = 1, 0 \leq q_j^t \leq 1 \quad \forall j = 0, 1, 2, \dots, Y_t$$

پس از حل معادلات فوق و تعیین استراتژی های مختلط، استراتژی های بهینه مرحله t بصورت زیر محاسبه می گردد:

$$(u_t^*, v_t^*) = \left(\sum_{i=0}^{X_t} i p_i^t, \sum_{j=0}^{Y_t} j q_j^t \right)$$

با جایگزینی استراتژی های مختلط محاسبه شده در توابع π_A و π_B توابع برگشت بهین مرحله t نیز بصورت زیر قابل حصول است:

$$f_t^*(X_t, Y_t) = \pi_A(P^t, Q^t)$$

$$g_t^*(X_t, Y_t) = \pi_B(P^t, Q^t)$$

(ح) مدل کنترل بهینه بازی های پیوسته

در این بخش متد تلفیقی مطرح شده در بخش های قبلی را به بازی های با مجموعه استراتژی های پیوسته توسعه می دهیم. بدین منظور از مدل کورنو استفاده می گردد که در آن دو شرکت به طور انحصاری بازار مربوط به یک محصول خاص را در اختیار دارند. هر یک از شرکتها بایستی نسبت به میزان تولید محصول مورد نظر در هر دوره از یک بازه n دوره ای تصمیم گیری نمایند بگونه ای که سود خویش را در مجموع این n دوره (با توجه به ظرفیت تولید n دوره ای X و Y) حداکثر نمایند.

مدل کورنو n دوره ای با در نظر گرفتن نرخ تنزیل

فرض کنید توابع سود شرکتها در دوره t بصورت زیر باشد:

$$L_t(u_t, v_t) = \delta^{t-1} \left((a - c)u_t - u_t^2 - u_t v_t \right)$$

$$K_t(u_t, v_t) = \delta^{t-1} \left((a - c)v_t - v_t^2 - u_t v_t \right)$$

متغیرها، شروط کمکی و معادلات تکراری در این بخش دقیقاً مطابق با تعاریف موجود در بخش قبلی می باشند با این تفاوت که فضای تصمیم از حالت گسسته به حالت پیوسته تغییر یافته است.

مرحله ۱. در مرحله اول از شروط کمکی زیر برای بدست آوردن مقادیر بهینه تابع هدف استفاده می نماییم :

$$f_1(X_1, Y_1) = \max_{u_1=X_1} L_1(u_1, v_1) = \max_{u_1=X_1} \{ (a - c)u_1 - bu_1^2 - bu_1 v_1 \}$$

$$g_1(X_1, Y_1) = \max_{v_1=Y_1} K_1(u_1, v_1) = \max_{v_1=Y_1} \{ (a - c)v_1 - bv_1^2 - bu_1 v_1 \}$$

مرحله ۲. مقادیر بهینه توابع هدف و استراتژی های بهینه در مرحله دوم را می توان از حل همزمان معادلات تکراری زیر بدست آورد:

$$\begin{aligned}
 f_2(X_2, Y_2) &= \max_{0 \leq u_2 \leq X_2} \{L_2(u_2, v_2) + f_1(X_2 - u_2, Y_2 - v_2)\} \\
 &= \max_{0 \leq u_2 \leq X_2} \{-(\delta + 1)u_2^2 + ((\delta - 1)(a - c) + 2X_2 + Y_2)u_2 - (\delta + 1)u_2v_2 + X_2v_2 \\
 &\quad + (a - c)X_2 - X_2^2 - X_2Y_2\} \\
 g_2(X_2, Y_2) &= \max_{0 \leq v_2 \leq Y_2} \{K_2(u_2, v_2) + g_1(X_2 - u_2, Y_2 - v_2)\} \\
 &= \max_{0 \leq v_2 \leq Y_2} \{-(\delta + 1)v_2^2 + ((\delta - 1)(a - c) + 2Y_2 + X_2)v_2 - (\delta + 1)u_2v_2 + Y_2u_2 + (a - c)Y_2 \\
 &\quad - Y_2^2 - X_2Y_2\}
 \end{aligned}$$

حال جهت تعیین استراتژی‌های بهینه از توابع فوق مشتق می‌گیریم. مشتقات جزئی مربوطه در زیر محاسبه گردیده است:

$$\begin{aligned}
 \frac{\partial f_2}{\partial u_2} &= -2(\delta + 1)u_2 + ((\delta - 1)(a - c) + 2X_2 + Y_2) - (\delta + 1)v_2 = 0 \\
 \frac{\partial g_2}{\partial v_2} &= -2(\delta + 1)v_2 + ((\delta - 1)(a - c) + 2Y_2 + X_2) - (\delta + 1)u_2 = 0
 \end{aligned}$$

حل همزمان دستگاه معادلات فوق، استراتژی‌های بهینه زیر را بدست می‌دهد:

$$\begin{aligned}
 u_2^* &= \frac{(\delta - 1)(a - c)}{3(\delta + 1)} + \frac{X_2}{\delta + 1} \\
 v_2^* &= \frac{(\delta - 1)(a - c)}{3(\delta + 1)} + \frac{Y_2}{\delta + 1}
 \end{aligned}$$

با جایگذاری مقادیر فوق در معادلات تکراری، مقادیر بهینه توابع هدف در مرحله دوم بصورت زیر محاسبه می‌گردند:

$$\begin{aligned}
 f_2(X_2, Y_2) &= \frac{(\delta - 1)^2(a - c)^2}{9(\delta + 1)} + \frac{2\delta(a - c)}{(\delta + 1)}X_2 - \frac{\delta}{(\delta + 1)}X_2^2 - \frac{\delta}{(\delta + 1)}X_2Y_2 \\
 g_2(X_2, Y_2) &= \frac{(\delta - 1)^2(a - c)^2}{9(\delta + 1)} + \frac{2\delta(a - c)}{(\delta + 1)}Y_2 - \frac{\delta}{(\delta + 1)}Y_2^2 - \frac{\delta}{(\delta + 1)}X_2Y_2
 \end{aligned}$$

ادعا می‌کنیم که اگر پارامترهای بازی بگونه‌ای باشند که u_i ها و v_i ها نامنفی شوند، آنگاه توابع برگشت بهین تعریف شده،

استراتژی‌های بهینه و توابع هدف بهینه زیر را در مرحله n ام بدست می‌دهند:

$$\begin{aligned}
 u_n^* &= \frac{1}{\sum_{i=0}^{n-1} \delta^i} \left(\frac{((\sum_{i=1}^{n-1} \delta^i) - (n - 1))(a - c)}{3} + X_n \right) \\
 v_n^* &= \frac{1}{\sum_{i=0}^{n-1} \delta^i} \left(\frac{((\sum_{i=1}^{n-1} \delta^i) - (n - 1))(a - c)}{3} + Y_n \right) \\
 f_n(X_n, Y_n) &= \frac{\delta^{n-1}}{\sum_{i=0}^{n-1} \delta^i} (n(a - c)X_n - X_n^2 - X_nY_n) + r_n(\delta) \\
 g_n(X_n, Y_n) &= \frac{\delta^{n-1}}{\sum_{i=0}^{n-1} \delta^i} (n(a - c)Y_n - Y_n^2 - X_nY_n) + r_n(\delta)
 \end{aligned}$$

که r_n در توابع فوق بصورت زیر تعریف می‌گردد:

$$r_n(\delta) = \frac{[\delta^{n-2}(\prod_{j=1}^{n-3}(\sum_{i=0}^j \delta^i))((\sum_{i=1}^{n-1} \delta^i) - (n - 1))^2 + (\sum_{i=0}^{n-1} \delta^i)r_{n-1}(\delta)](a - c)^2}{9(\prod_{j=1}^{n-1}(\sum_{i=0}^j \delta^i))}$$

اثبات ادعای فوق در پیوست ۱ آمده است.

حال جهت بدست آوردن دنباله تصمیم بهینه و مسیر بهینه، از انتها به ابتدا حرکت نموده و با استفاده از دینامیک مسأله، استراتژی‌های بهینه در هر مرحله را بدست می‌آوریم. می‌دانیم که در مرحله n ام $X_n = X$ و $Y_n = Y$ می‌باشد لذا با جایگذاری آنها، متغیرهای تصمیم بهینه در مرحله n ام یعنی u_n^* و v_n^* بدست می‌آیند. سپس با استفاده از دینامیک مسأله یعنی $Y_{n-1} = Y_n - v_n$ و $X_{n-1} = X_n - u_n$ ، متغیرهای حالت مرحله قبل یعنی $n - 1$ محاسبه می‌گردند. با تعیین متغیرهای حالت مرحله $n - 1$ و جایگذاری آنها، متغیرهای تصمیم بهینه در مرحله مورد نظر بدست می‌آیند. به همین ترتیب پیش رفته و متغیرهای تصمیم بهینه در هر مرحله حاصل می‌گردند.

نکته جالب توجهی که از نتایج فوق حاصل می‌گردد این است که در صورت عدم در نظر گرفتن نرخ تنزیل (یعنی در شرایطی که $\delta = 1$ باشد) کل ظرفیت تولید به صورت مساوی در n دوره جهت عرضه پیشنهاد می‌گردد. یعنی روش بکارگرفته شده، توابع هدف بهینه و استراتژی‌های بهینه زیر را در مرحله n ام بدست می‌دهد:

$$\begin{aligned} u_1^* &= u_2^* = \dots = u_n^* = \frac{X}{n} \\ v_1^* &= v_2^* = \dots = v_n^* = \frac{Y}{n} \\ f_n(X, Y) &= (a - c)X - \frac{b}{n}X^2 - \frac{b}{n}XY \\ g_n(X, Y) &= (a - c)Y - \frac{b}{n}Y^2 - \frac{b}{n}XY \end{aligned}$$

۳- نتایج و بحث

تحلیل بازار رقابتی و کمک به تصمیم‌گیری بنگاه‌های اقتصادی یکی از مهم‌ترین مسائلی است که در دنیای امروزی پژوهشگران اقتصادی با آن روبه‌رو هستند. به دلیل تأثیر این تصمیم‌ها بر سودآوری و بقای سازمان‌ها در بازار رقابت، مدیران بنگاه‌های اقتصادی از اهمیت این تصمیم‌ها آگاه هستند. آن‌ها باید با کنترل و تعیین مقدار بهینه این متغیرها بتوانند در هر دوره بهترین تصمیم را به منظور بقاء در عرصه رقابت بگیرند. تکنیک تلفیقی نظریه بازی و کنترل بهینه ارائه شده در این پژوهش، راهبردهای بهینه را به بازیکنان ارائه می‌نماید. در این پژوهش، یک مدل تصمیم‌گیری یکپارچه ارائه شده است که دو چارچوب تحلیلی اصلی یعنی کنترل بهینه گسسته‌زمان و نظریه بازی‌های غیرهمکارانه را با یکدیگر ترکیب می‌کند. این مدل به شرکت‌های فعال در بازارهای انحصاری دوقطبی کمک می‌کند تا تصمیم‌های تولید چنددوره‌ای بهینه‌ای اتخاذ کنند؛ به گونه‌ای که هم اهداف خود را دنبال کرده و هم واکنش‌های استراتژیک رقیب را طی زمان در نظر بگیرند. نوآوری اصلی این پژوهش در ادغام چارچوب کنترل بهینه گسسته با مدل‌سازی نظریه بازی‌ها نهفته است. در حالی که بیشتر مطالعات پیشین این دو رویکرد را به صورت مجزا بررسی کرده‌اند، مدل حاضر تلاش دارد این دو دیدگاه را در قالب یک مدل پویا و کاربردی تلفیق کند. مدل‌های سنتی نظریه بازی‌ها اغلب جنبه پویایی زمان را نادیده می‌گیرند و از سوی دیگر، مدل‌های کنترل بهینه کلاسیک، رفتار راهبردی رقبا را صریحاً وارد محاسبات خود نمی‌کنند. این پژوهش این خلأ را با مدلی که رقابت دو شرکت را در قالب یک بازی چندمرحله‌ای گسسته در طول زمان مدل‌سازی می‌کند، برطرف کرده است. در این بازی، هر شرکت در هر دوره زمانی تصمیم بهینه خود را می‌گیرد و در عین حال، به رفتار احتمالی رقیب نیز واکنش نشان می‌دهد. در چارچوب کنترل بهینه گسسته، تکامل متغیرهای حالت مانند منابع باقی‌مانده، موجودی کالا یا تولید تجمعی، در طول زمان مدل‌سازی می‌شود. تصمیم‌گیری‌ها در فواصل زمانی مشخص (مثلاً فصلی یا سالانه) صورت می‌گیرند و اثر آن‌ها بر دوره‌های آتی صریحاً در نظر گرفته می‌شود. نکته مهم آن است که زمان به عنوان یک پارامتر مستقل و غیرقابل کنترل در نظر گرفته می‌شود، برخلاف برخی مدل‌های پیشین که نقش زمان را ساده‌سازی یا حذف می‌کردند. این امر امکان مدل‌سازی واقع‌گرایانه‌تری از تصمیم‌های تولیدی را فراهم می‌کند، به ویژه در صنایعی که سرعت واکنش و زمان‌بندی اهمیت بالایی دارد.

در سوی دیگر، نظریه بازی‌ها برای مدل‌سازی وابستگی استراتژیک میان دو شرکت به کار می‌رود. در هر مرحله، هر شرکت تلاش می‌کند سود خود را بیشینه کند، در حالی که واکنش احتمالی رقیب را نیز پیش‌بینی می‌کند. این تعامل در قالب یک بازی غیرهمکارانه با مجموعه مشخصی از استراتژی‌ها و تابع‌های سوددهی تعریف می‌شود. برای تحلیل پایداری نتایج، از مفهوم تعادل نش استفاده شده و در برخی حالت‌ها، پویایی استکلبرگ (رهبر-پیرو) نیز بررسی شده است. برای حل مدل تلفیقی، از الگوریتمی بر پایه استقرا معکوس که از اصول برنامه‌ریزی پویا الهام گرفته شده، بهره گرفته شده است. در این رویکرد، تصمیم بهینه هر مرحله بر مبنای سود فعلی و سود مورد انتظار در دوره‌های آتی محاسبه می‌شود. الگوریتم از دوره پایانی به سمت دوره آغازین حرکت می‌کند و به صورت بازگشتی، تابع ارزش و استراتژی بهینه را در هر گام به‌روزرسانی می‌کند. این ویژگی باعث می‌شود هر تصمیم با لحاظ اثر آن در آینده اتخاذ شود. برای اعتبارسنجی مدل و نشان دادن کاربرد عملی آن، شبیه‌سازی‌های عددی با پارامترهای واقعی انجام شده است. در این شبیه‌سازی‌ها، عواملی مانند هزینه‌های ثابت و متغیر، ظرفیت تولید، تقاضای بازار،

ضریب تنزیل، و کشش قیمتی وارد مدل شده‌اند. نتایج نشان می‌دهد مدل پیشنهادی می‌تواند مقدار بهینه تولید هر شرکت را در هر دوره زمانی به خوبی محاسبه کند، در حالی که پویایی رقابت و سودآوری بلندمدت را نیز در نظر دارد. همچنین، نتایج مدل با مدل‌هایی که تنها از یک چارچوب (نظریه بازی) استفاده می‌کنند، مقایسه شده‌اند. این مقایسه‌ها نشان می‌دهد مدل ترکیبی عملکرد بهتری دارد و تصمیم‌هایی متعادل‌تر و راهبردی‌تر تولید می‌کند. برخلاف مدل‌های ایستا که بر تصمیم‌های کوتاه‌مدت تمرکز دارند، یا مدل‌های کنترلی که رفتار رقیب را نادیده می‌گیرند، این مدل هر دو جنبه ساختار زمانی و محیط رقابتی را به طور همزمان در نظر می‌گیرد. یکی از یافته‌های کلیدی مدل آن است که استراتژی‌های وابسته به زمان تفاوت چشم‌گیری با استراتژی‌های ایستا دارند. به عنوان مثال، یک شرکت ممکن است در دوره‌های ابتدایی، با کاهش سود کوتاه‌مدت، به طور تهاجمی به دنبال کسب سهم بازار باشد تا در دوره‌های آتی بر بازار تسلط یابد. یا بالعکس، ممکن است گسترش تولید خود را به تأخیر بیندازد تا از محدودیت ظرفیت رقیب بهره‌برداری کند. این رفتارهای ظریف، تنها در مدل‌هایی با ساختار بازگشتی و تعاملی قابل شناسایی هستند. به طور خلاصه، این پژوهش چارچوبی دقیق، یکپارچه و منعطف ارائه می‌دهد که شرکت‌ها را قادر می‌سازد در بازارهای انحصاری پویا و رقابتی دوقطبی، تصمیم‌هایی آگاهانه، بهینه و راهبردی اتخاذ کنند. با تلفیق اصول نظریه بازی‌ها و کنترل بهینه، مدل پیشنهادی هم از لحاظ نظری غنی است و هم در عمل کاربردی بوده و می‌تواند مبنای تصمیم‌سازی برای پژوهشگران، مدیران و سیاست‌گذاران صنعتی باشد.

هرچند این مقاله گامی به جلو برای عملیاتی شدن مسأله مورد بحث است، اما هنوز جنبه‌های مختلفی برای کامل شدن مسأله مورد نیاز است. لذا می‌توان برای کارهای آینده و جهت روند آتی پژوهش موارد زیر را پیشنهاد نمود:

۱. مدل ارائه شده تنها جوابگوی حل بازی‌های دونفره از قبیل مدل‌های انحصار دوقطبی می‌باشد؛ اما اینکه چگونه می‌توان بازی-های n نفره از قبیل مدل‌های انحصار چند جانبه را با این تکنیک حل نمود هنوز بدون پاسخ است. ارائه راهکار برای حل این مدل‌ها می‌تواند راهگشای مشکلات بازارهای مهمی از قبیل بازار جهانی نفت ایک گردد.

۲. ایده‌ای که تکنیک برنامه‌ریزی پویای مطروحه جهت بهینه‌سازی و پیشروی حل مسائل انحصار دوقطبی در هر مرحله در نظر می‌گرفت، ایده تعادل نش و استکلبرگ بود. حال این سؤال مطرح می‌گردد که آیا ایده تعادل دیگری غیر از تعادل نش برای ماکزیمم نمودن سود بازیکنان وجود دارد یا خیر؟ به طور مثال آیا می‌توان با تعریف استراتژی‌های ماشه مناسب به سود بیشتری نسبت به حالت کنونی دست یافت؟ یا اینکه به جای رویکرد تقابلی و رقابتی، رویکرد تعاملی و همکاری را در پیش گرفته و با استفاده از مدل‌های موجود در نظریه بازی‌های مشارکتی، مزایای حاصل از همکاری را در قالب نتایج تحلیلی، ارائه نمود.

۳. مدل‌های ارائه شده در این مقاله، میزان عرضه کالا را به عنوان متغیرهای تصمیم یا استراتژی‌های مناسب در مدل‌های انحصار دوجانبه پویا، مورد بررسی قرار می‌داد لذا مدل‌سازی و حل مسأله رقابت در قیمت‌گذاری بین دو بنگاه تولیدی در قالب توسعه مدل برتراند با توجه به مدل و الگوریتم ارائه شده در مطالعه حاضر، می‌تواند مورد دیگری جهت ادامه مطالعه باشد. همچنین در نظر گرفتن رقابت در قیمت و میزان عرضه محصول بطور همزمان نیز که بر پیچیدگی مدل می‌افزاید، می‌تواند موضوع جالب توجه و کاربردی دیگری جهت مطالعات آتی باشد.

۴. یکی از فرض‌هایی که در این تحقیق در نظر گرفته شده است، وجود اطلاعات کامل و متقارن برای بازیکنان است. در بازی‌های با اطلاعات کامل، تمامی بازیکنان، ترجیحات بازیکنان دیگر را می‌دانند، درحالیکه در بازی با اطلاعات غیر کامل، این دانش بطور کلی وجود ندارد. بازی‌های با اطلاعات غیر کامل، بسیاری از تعاملات راهبردی در دنیای واقعی کسب و کار را منعکس می‌کنند. بنگاه‌ها اغلب اطلاعات خصوصی مختلفی دارند که باعث می‌شود عدم تقارن اطلاعات بصورت ذاتی در بازار وجود داشته باشد. لذا بررسی پویایی رقابت و همکاری در این حالت نیز از موضوعات مهم دیگر، جهت ادامه روند پژوهش خواهد بود.

۴- منابع

- Askar, S. S., & Alnowibet, K. (2016). Cooperation versus noncooperation: Cournot duopolistic game based on delay and time-dependent parameters. *Chaos, Solitons and Fractals*, 91, 580–584. <https://doi.org/10.1016/j.chaos.2016.07.012>
- Aflaki Sam, M. (2006). *Decision making based on game theory* [Master's thesis, Sharif University of Technology].

- Barquin, J., Garcia-Gonzalez, J., & Roman, J. (2000, September 25–28). *Water value in competitive markets: Dynamic programming and game theory* [Conference presentation]. 6th Probabilistic Methods Applied to Power Systems Conference, Madeira, Portugal.
- Basiri, R., Abedian, M., Aghasi, S., & Dashtaali, Z. (2024). Presenting a model to determine the equilibrium in an electricity oligopoly with strategic investment decisions: A case study. *Journal of Applied Theories of Economics*, 11(2), 63–94.
- Bellman, R., & Kalaba, R. (1965). *Dynamic programming and modern control theory*. Academic Press.
- Boltyanskii, V. G. (1973). *Optimal control of discrete systems*. Nauka.
- Cournot, A. A. (1971). *Researches into the mathematical principles of the theory of wealth* (N. T. Bacon, Trans.). Augustus M. Kelley. (Original work published 1838)
- Dori, M., Jafari, M., & Chaharsoghi, K. (2021). Competition between two supply chains with demand dependent on price and lead time: A games theory approach. *Modern Research in Decision Making*, 6(3), 23–49.
- Faraji Dizaji, S. (2010). *Microeconomic theory*. Iran University of Economic Science.
- Farzaneh Tatari, F., & Naghibi-Sistani, M. B. (2015). Distributed optimal control of nonlinear differential graphical games based on reinforcement learning. *Journal of Control*, 8(4), 15–30.
- Garay, J., Csiszar, V., & Mori, T. F. (2017). Evolutionary stability for matrix games under time constraints. *Journal of Theoretical Biology*, 415, 1–12. <https://doi.org/10.1016/j.jtbi.2016.11.023>
- Hamdy, A. T. (1976). *Operations research: An introduction*. Macmillan.
- Itza-Ortiz, B. A., Villafuerte-Segura, R., & Alvarado-Santos, E. (2020). Delay-dependent and delay-independent stability of Cournot duopoly model with tax evasion and time-delay. *Communications in Nonlinear Science and Numerical Simulation*, 87, Article 105274. <https://doi.org/10.1016/j.cnsns.2020.105274>
- Jozsef, G., & Varga, T. (2024). Dynamically unstable ESS in matrix games under time constraints. *Journal of Mathematical Biology*, 88(1), 1–32. <https://doi.org/10.1007/s00285-023-02020-5>
- Kelly, A. (2003). *Decision making using game theory: An introduction for managers*. Cambridge University Press.
- Krabs, W., & Pickl, S. (2003). *Analysis, controllability and optimization of time-discrete systems and dynamical games* (Lecture Notes in Economics and Mathematical Systems, Vol. 529). Springer-Verlag. <https://doi.org/10.1007/978-3-642-18927-5>
- Krivan, V., & Cressman, R. (2017). Interaction times change evolutionary outcomes: Two-player matrix games. *Journal of Theoretical Biology*, 416, 199–207. <https://doi.org/10.1016/j.jtbi.2017.01.010>
- Larson, R. E., & Casti, J. L. (1978). *Principles of dynamic programming: Part I, basic analytic and computational methods*. Marcel Dekker.
- Lozovanu, D., & Pickl, S. (2006). An approach for an algorithmic solution of discrete optimal control problems and their game-theoretical extension. *Central European Journal of Operations Research*, 14(4), 357–375. <https://doi.org/10.1007/s10100-006-0011-x>
- Ma, J., Wang, Z., & Dai, Y. (2020). A supply chain network economic model with time-based competition. *European Journal of Operational Research*, 280(3), 889–908. <https://doi.org/10.1016/j.ejor.2019.07.066>
- Mahmoudinia, D., Dallali Esfahani, R., Engwerda, J., & Dastjerdi, R. (2017). Game theory and its role in determining optimal policies and strategic interaction between fiscal and monetary policymakers (Application of differential game theory and Stackelberg games). *Quarterly Journal of Applied Economics Studies Iran*, 5(18), 1–34.
- Mark, B., & Krivan, V. (2020). Two-strategy games with time constraints on regular

- graphs. *Journal of Theoretical Biology*, 506, Article 110426. <https://doi.org/10.1016/j.jtbi.2020.110426>
- McMillan, J. (1996). *Games, strategies, and managers: How managers can use game theory to make better business decisions*. Oxford University Press.
- Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1), 48–49. <https://doi.org/10.1073/pnas.36.1.48>
- Neumann, J. von, & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.
- Pekgün, P., Griffin, P. M., & Keskinocak, P. (2017). Centralized versus decentralized competition for price and lead-time sensitive demand. *Decision Sciences*, 48(6), 1097–1130. <https://doi.org/10.1111/deci.12264>
- Raoufinia, M., Baradaran, V., & Shahrjerdi, R. (2018). A dynamic game theory model for analyzing competition in the oligopoly markets. *Journal of Industrial Management Perspective*, 8(2), 63–94.
- Shahikitash, M., & Mohammadzadeh, A. (2014). Measurement of conjectural variations elasticity in an oligopoly structure (Using Iwata Approach for Iranians food and beverage industry). *Journal of Agricultural Economics and Development*, 28(3), 272–281. <https://doi.org/10.22067/jead2.v0i0.34292>
- Straffin, P. D. (1993). *Game theory and strategy*. Mathematical Association of America.
- Vamvoudakis, K. G., & Lewis, F. L. (2011). Multi-player non-zero-sum games: Online adaptive learning solution of coupled Hamilton–Jacobi equations. *Automatica*, 47(8), 1556–1559. <https://doi.org/10.1016/j.automatica.2011.03.005>
- Veraguas, J. B., & Zhang, X. (2023). Dynamic Cournot–Nash equilibrium: The non-potential case. *Mathematics and Financial Economics*, 17, 153–174. <https://doi.org/10.1007/s11579-022-00323-7>
- Wu, D., Baron, O., & Berman, O. (2009). Bargaining in competing supply chains with uncertainty. *European Journal of Operational Research*, 197(2), 548–556. <https://doi.org/10.1016/j.ejor.2008.07.017>
- Zhang, H., Cui, L., & Luo, Y. (2013). Near-optimal control for nonzero-sum differential games of continuous-time nonlinear systems using single-network ADP. *IEEE Transactions on Cybernetics*, 43(1), 206–216. <https://doi.org/10.1109/TSMCB.2012.2203335>

EXTENDED ABSTRACT

Development of an Optimal Control Model in Decision Making of Duopoly using Nash and Stackelberg Equilibrium

Syed Jalal Ziaei (Corresponding Author)

Industrial Engineering Department, Sharif University of Technology, Tehran, Iran.

Email: ziaei@yahoo.com

Rasoul haji

Professor of Industrial Engineering Department, Sharif University of Technology, Tehran, Iran

Introduction

In today's highly competitive and complex business environment, strategic decision-making by firms depends not only on internal factors but also on the actions of competitors, the characteristics of the market structure, and dynamic environmental changes over time. This dependency becomes especially critical in markets characterized by duopoly structures, where two dominant firms compete, and each firm's decisions directly influence the profits and losses of the other. Under such conditions, the strategic interactions are highly interdependent, making optimal decision-making a complex challenge. This paper aims to develop a comprehensive and innovative analytical framework for firms to make optimal decisions in dynamic and adversarial environments by integrating two powerful theoretical approaches: discrete optimal control and game theory.

The study begins by introducing the theoretical framework of discrete optimal control, which facilitates the modeling of systems where decision-making occurs over multiple discrete time periods. In such systems, each decision not only affects immediate outcomes but also has consequences for future states, linking past, present, and future periods in a dynamic process. A key contribution of this research is the explicit treatment of time as an independent and influential parameter in the decision-making process, contrasting with traditional models that often ignore or treat time indirectly. By incorporating time explicitly, the model gains a greater capacity to capture the evolving nature of strategic interactions and the temporal dependencies inherent in competitive markets.

Methodology

To examine the problem in a realistic context, the study focuses on the duopolistic market structure, a common and significant form of market competition where two firms produce similar or substitutable products. Each firm seeks to determine its production and supply strategy over a finite number of time periods, considering production capacity limitations and the dynamic behavior of the rival. The objective is to maximize cumulative profits over these periods while taking into account the competitor's potential responses. Recognizing the strategic interdependence, the proposed model integrates concepts from both Nash equilibrium, which assumes simultaneous decision-making, and Stackelberg equilibrium, which models sequential moves and leadership dynamics in competitive settings.

To solve the resulting complex dynamic optimization problem, the paper develops an algorithm that combines dynamic programming techniques with game-theoretical analysis. This approach allows for sequential optimization where, at each discrete time step, firms assess the current market conditions and forecast competitors' strategies to determine their own optimal supply decisions. The novelty of this approach lies in treating time as an external parameter that influences the system dynamics but is not directly controlled by the players. This treatment enhances the realism and applicability of the model to practical business situations, where time

constraints and external temporal factors often affect firm behavior.

The practical relevance of the model is demonstrated through two numerical examples in which firms must decide on the quantity of product to supply in each period. These examples validate the proposed framework by showing that the combined use of game theory and optimal control leads to more accurate and strategically sound decisions compared to approaches relying solely on either theory. This integration reflects the complexity of competitive decision-making in real markets and provides firms with a more sophisticated tool for planning and adapting their strategies over time.

In addition to the model development and numerical validation, the paper provides a comprehensive review of the theoretical literature related to game theory and optimal control. It highlights a notable gap in prior research: while the role of time in strategic games has been acknowledged, few studies have treated time as an independent and influential external parameter affecting game outcomes directly. Many existing models treat time as a controlled variable within players' strategies or as part of the decision process itself. By contrast, this research models time as an exogenous factor, external to the players' control but critical in shaping decision outcomes, offering a fresh perspective on temporal dynamics in competitive interactions.

Results Discussion

Beyond the conceptual innovation, this paper outlines several promising avenues for future research. These include extending the model to multiplayer games involving more than two firms, incorporating both price and quantity decision variables for a richer strategic space, and investigating games with incomplete information and asymmetric knowledge among competitors. Furthermore, the possibility of applying alternative equilibrium concepts, such as trigger strategies or cooperation-oriented equilibria, is discussed as a way to capture a wider range of strategic behaviors and market conditions. There is also the potential for the model to be developed and applied to real industries such as the OPEC oil market, the automotive industry, and other industries with high market concentration and strategic dependencies between a limited number of companies.

In summary, this study proposes a robust and flexible modeling framework that allows firms to navigate the complexities of dynamic, competitive decision-making in duopolistic markets. By unifying game-theoretic reasoning with optimal control principles, the model provides both theoretical richness and practical value. It offers a comprehensive tool for academics, managers, and policymakers interested in modeling, analyzing, and optimizing strategic interactions in dynamic industrial environments.

Through analytical modeling and numerical simulations, the study yielded several key findings. The proposed integrated framework demonstrates a dynamic strategic balance, capturing how firms continuously adapt their production in response to competitors' actions and market feedback, resulting in evolving rather than static strategies. This approach leads to improved long-term profitability compared to static models, as it leverages strategic anticipation and foresight. Furthermore, it enables efficient resource allocation, guiding firms to avoid overproduction and ensure sustainable operations throughout the planning horizon. The model also provides a mechanism for strategic market share management, allowing firms to make calculated trade-offs between short-term gains and long-term market positioning. A central comparative advantage of this framework is its integration of control and game theory, unlike studies that treat them in isolation. Optimal control ensures temporal consistency in decision-making across the planning horizon, while game theory rigorously accounts for strategic interactions. This synthesis allows for more robust and realistic modeling of competitive dynamics over time.