

تحلیل نظریه بازی تکاملی ایران و عربستان در چارچوب الگوریتم ژنتیک^۱

سمانه خاتمی*، علیرضا شکیبائی**

تاریخ دریافت: ۹۵/۱۲/۲۳ تاریخ پذیرش: ۹۶/۰۴/۱۸

چکیده

هدف این مقاله ارائه مدلی جدید از جستجوی استراتژی‌های بهینه در بازی معمای زندانی تکراری با استفاده از الگوریتم ژنتیک است. بدین منظور با شبیه‌سازی رقابت بین ایران و عربستان در ائتلاف اوپک نفتی، از ۱۲ نوع استراتژی مطرح در بازی معمای زندانی تکراری طی ۲۰ اجرای الگوریتم ژنتیک به منظور حداکثرسازی امتیازات فردی بازیکن و نیز حداقل‌سازی امتیاز برابری رقیب استفاده شده است. نتایج نشان داد استراتژی "عمل متقابل" حائز بالاترین بازدهی متوسط در هر دو رقابت بوده و در رتبه‌های بعدی استراتژی‌های "اکثریت موافق"، "ماشه" و "عمل متقابل پس از دو بار نقض همکاری رقیب" جای گرفته‌اند. استراتژی "همواره عدم همکاری" نیز در رقابت‌ها با کمترین بازدهی به‌عنوان ناکارترین استراتژی شناخته شده است.

طبقه‌بندی JEL: C73, C70, C63

واژگان کلیدی: بازی، محاسبات تکاملی، الگوریتم ژنتیک، معمای زندانی تکراری، اوپک.

^۱ مقاله حاضر برگرفته از رساله دکتری سمانه خاتمی به راهنمایی دکتر علیرضا شکیبائی در دانشگاه شهید باهنر کرمان است.

* دانشجوی دکتری علوم اقتصادی، دانشگاه شهید باهنر کرمان (نویسنده مسئول)، ایران، پست الکترونیکی:

khatamisamaneh@gmail.com

ashakibae@yahoo.com

** دانشیار گروه اقتصاد دانشگاه شهید باهنر کرمان، ایران، پست الکترونیکی:

۱. مقدمه

نظریه بازی مجموعه‌ای از ابزارهای تحلیلی است که به فهم پدیده‌های به وجود آمده در هنگام برهم‌کنش میان تصمیم‌گیرندگان کمک می‌نماید (سامتی و همکاران، ۱۳۹۰). کاربرد نظریه بازی^۱ (GT) در اقتصاد از سال ۱۹۴۴ با انتشار کتاب "نظریه بازی‌ها و رفتار اقتصادی" وون نیومن و مورگنسترن^۲ آغاز و با سرعتی شگفت‌آور در ارائه انواع مشتقه آن رواج یافت. از آن جمله می‌توان نظریه بازی‌های تکاملی^۳ (EGT) را برشمرد که در آن فرض عقلانیت کامل بازیکنان در GT کلاسیک (به معنای برخورداری از دانش کامل درباره محیط و جدول بازدهی‌ها)، تعدیل می‌گردد.

در این راستا، مطالعه "تکامل و نظریه بازی‌ها" جان مینارد اسمیت^۴ به منزله پایه و اساس مطالعات EGT شناخته شده است. در این مطالعه اسمیت، بر اساس شرایط بیولوژیکی لحاظ شده، قضاوت در مورد عقلانی بودن یا نبودن انتخاب‌های بازیکنان را غیرممکن عنوان کرده و لذا بهینه‌سازی رفتار فردی بر اساس دانش محدود عوامل مطرح می‌گردد (اسمیت، ۱۹۸۲).

همچنین اسمیت مفاهیم اساسی EGT از جمله استراتژی پایدار تکاملی^۵ (ESS) را معرفی می‌نماید. چنین استراتژی در مفهوم تعادلی به معنای پایداری در برابر جهش‌های مهاجم بالقوه است. در واقع، اسمیت با تفکیک نظریه بازی از کاربردهای اقتصادی، از آن به عنوان ابزاری برای تحلیل تکامل بیولوژیکی بهره گرفته و از این‌رو، با معرفی نظریه بازی تکاملی، برازنده‌ترین ابزارها برای مدل‌بندی پویایی‌های تعاملات استراتژیک را فراهم نموده است.

از سوی دیگر، نوع تصمیم‌گیری عوامل با عقلانیت محدود در بازی‌های تکاملی بر اساس «فرایند یادگیری» تدریجی بازیکنان صورت می‌گیرد. فرایند یادگیری عوامل در بازی‌های تکاملی نیز با استفاده از الگوریتم‌های تکاملی^۶ (EA) محقق می‌گردد. الگوریتم‌های تکاملی بنا به تعریف آلکمد^۷ (۲۰۰۴) عبارتند از: تکنیک‌های مبتنی بر علوم زیستی که با استفاده از مفهوم "بقای اصلح" برای تکامل رفتار هر عامل، منجر به انطباق بهتر رفتار آن عامل با محیط

^۱ Game Theory

^۲ Von Neumann & Morgenstern

^۳ Evolutionary Game Theory

^۴ John Maynard Smith

^۵ Evolutionarily Stable Strategy

^۶ Evolutionary Algorithms

^۷ Alkemade

(برای مثال یک بازار خاص) می‌شوند. بنابراین، الگوریتم‌های تکاملی، روش نوینی را برای مدل‌بندی یادگیری و تصمیم‌گیری عوامل با عقلانیت محدود ارائه می‌دهند. در مجموع، هدف یک الگوریتم تکاملی بهینه‌سازی تابع هدف برازندگی است (دوبوز و همکاران^۱، ۲۰۱۰). از جمله الگوریتم‌های تکاملی می‌توان به الگوریتم ژنتیک^۲ (GA) اشاره کرد که اغلب در اقتصاد به منظور توصیف چارچوبی مناسب و دقیق از یادگیری اجتماعی مورد استفاده قرار می‌گیرد. برخی مطالعات برجسته در مورد استفاده از الگوریتم ژنتیک برای فرمول‌بندی دقیق رفتار عوامل در بازی‌های اقتصادی عبارتند از: مطالعه آکسلرود^۳ در سال ۱۹۸۱ و نیز مطالعه مارکس^۴ در سال ۱۹۹۲. همچنین الگوریتم‌های ژنتیک در مورد مسائل داخلی و چالش‌برانگیز اقتصادی نیز به کار گرفته شده و به لحاظ خواص پویایی و تصادفی بودن آن‌ها در اقتصاد مورد توجه قرار گرفته‌اند.

از این‌رو، رویکرد مطالعه حاضر، استفاده از الگوریتم ژنتیک به عنوان یک الگوریتم تکاملی برای حل یک بازی تکاملی است. هدف اصلی در این پژوهش با توجه به ماهیت چندبعدی بازارهای جهانی نفت و محیط اقتصادی پیچیده آن و عدم امکان دست‌یابی به نتایج تجربی قطعی و معین، بهره‌گیری از روشی نوین و کاربردی است که بر مبنای فرایندهای "یادگیری" و "جهش" رفتار عوامل باشد. لذا در این خصوص به ارائه ابزاری روش‌شناختی با ترکیب EGT و الگوریتم ژنتیک پرداخته می‌شود.

این نوع تحلیل بازار انرژی متفاوت از تکنیک‌های مدل‌سازی اقتصادی متعارف مبتنی بر فرض عقلانیت کامل عوامل می‌باشد. در این راستا به منظور جستجوی استراتژی‌های بهینه اعضای اوپک، با لحاظ دو کشور "ایران" و "عربستان" به عنوان نمایندگان دو گروه عوامل ناهمگن اوپک و با استفاده از رویکرد تکاملی، بازی معمای زندانی تکراری بین این دو بازیکن در محیطی شبیه‌سازی شده مطرح می‌شود. همچنین ۱۲ نوع استراتژی تصادفی برای هر یک از بازیکنان در نظر گرفته شده که طی ۲۰ اجرای الگوریتم ژنتیک با هدف دستیابی به استراتژی بهینه حداکثرکننده امتیاز فردی و نیز حداقل‌کننده امتیاز رقیب در بازی معمای زندانی تکراری^۵

¹ Duboz Et Al

² Genetic Algorithm

³ Axelrod

⁴ Marks

⁵ Iterated Prisoner's Dilemma

(IPD) توسط اعضاء انتخاب می‌شوند. طبق نتایج گزارش شده با استفاده از الگوریتم تکاملی، استراتژی TFT با بالاترین بازدهی به عنوان استراتژی بهینه در مقایسه با دیگر استراتژی‌ها در هر دو بازی برای هر یک از بازیکنان شناخته شده است.

از آنجا که با حل معمای زندانی در قالب کلاسیک، عدم همکاری به عنوان تعادل پایدار شناخته شده و با یکبار انجام بازی ظهور و تکامل همکاری غیرممکن به نظر می‌رسد. لکن با تکرار بازی در راندهای مختلف و لذا با تکیه بر حافظه بازیکنان در به‌خاطر سپردن حرکت‌های رقیب در دوره‌های پیشین و عکس‌العمل به آن‌ها با عمل مقابله به‌مثل، استراتژی همکاری می‌تواند به عنوان تعادل پایدار تکاملی شناخته شود.

بنابراین، چارچوب مطالعه حاضر مبنی بر مروری بر ادبیات، بررسی مبانی نظری ساختار الگوریتم ژنتیک و ارائه مفاهیم پویایی و تکاملی بودن آن می‌باشد. در ادامه، روش تحقیق بر اساس مبانی مطالعه بهینه‌سازی تکاملی آکسلرود (۱۹۸۱) ارائه شده و نهایتاً مدل‌سازی IPD با استفاده از الگوریتم تکاملی و نتایج امتیازات استراتژی‌های ۱۲ گانه طی رقابت شبیه‌سازی شده دو کشور توسط الگوریتم ژنتیک ارائه می‌شوند.

۲. مروری بر ادبیات

۲-۱. اوپک؛ ساختار و سیاست‌ها

نفت به عنوان منبع انرژی اصلی و برخوردار از صرفه‌های بزرگ اقتصادی ناشی از پایین بودن هزینه‌ها، در مقایسه با سوخت‌های دیگر از میزان انرژی بالاتری نیز برخوردار است. به عبارتی دارای انرژی حدود ۵۰ درصد بیشتر از زغال‌سنگ بر پایه وزن و ۱۷۰ برابر بیشتر از گاز طبیعی بر مبنای حجم می‌باشد. لذا نفت خام و محصولات پالایش شده بزرگ‌ترین بخش را در تجارت بین‌الملل چه بر اساس ارزش یا حجم تشکیل می‌دهند. از این‌رو، تجارت نفت به‌عنوان تجارتي استراتژیک، بین‌المللی و بسیار حائز اهمیت شناخته شده است (جمشیدی رودباری، ۱۳۸۷).

ظهور اوپک به عنوان قدرتی فزاینده در بازارهای جهانی نفت، به درستی بیانگر این واقعیت است که کشورهای عضو اوپک سهم بزرگی از صادرات جهانی نفت را عهده‌دار بوده و مالک

بخش عمده ذخایر نفتی موجود در زمین می‌باشند؛ به طوری که بیش از ۷۵ درصد از ذخایر نفت جهان در ۱۲ کشور عضو اوپک متمرکز شده و لذا تقریباً تمامی کشورها به اوپک وابسته‌اند. حتی می‌توان گفت خود کشورهای عضو اوپک نیز به بقای این سازمان وابسته‌اند؛ زیرا بیشتر اعضاء بین ۸۰ تا ۹۹ درصد به درآمدهای ارزی ناشی از فروش نفت وابستگی داشته و این توانایی سازمان است که قیمت‌ها را حفظ نموده تا حداکثر درآمد را برای اعضایش به ارمغان آورد (عبدلی و ناخدا، ۱۳۸۸).

با وجود این، تنوع کشورهای عضو اوپک و تفاوت دیدگاه‌های سیاسی، اقتصادی، اجتماعی و فرهنگی، به عنوان منبع ضعف بالقوه‌ای در ائتلاف محسوب می‌شود؛ زیرا اعضای اوپک برحسب جمعیت، نیازهای مالی و سرمایه‌گذاری، ذخایر نفتی، ظرفیت تولید نفت و درآمد سرانه، تفاوت‌های چشم‌گیری با یکدیگر دارند. بنابراین، درون این سازمان دودستگی بین کشورهای عضو قابل بررسی است (گریفین و ویلهابر^۱، ۱۹۹۴).

گروه کشورهای اوپک بر اساس تفاوت‌های ماهوی به دو گروه قابل تقسیم هستند. گروه اول به رهبری عربستان سعودی شامل کشورهای کویت، قطر، امارات متحده عربی و لیبی است که از درآمد سرانه بالا، ذخایر اثبات شده سرانه قابل توجه برخوردار هستند (کشورهای سازگار با آینده). این کشورها همان‌گونه که تاریخ نشان می‌دهد طرفدار افزایش عرضه کل نفت در تلاش برای تعدیل قیمت‌ها و حفظ تقاضای بلندمدت بوده‌اند.

گروه دوم نیز شامل کشورهای ایران، نیجریه، عراق، ونزوئلا، آنگولا و الجزایر (کشورهای سازگار با حال) خواستار این هستند که اوپک کل تولیداتش را به نفع افزایش سریع قیمت‌ها محدود نماید. این کشورها دارای درآمد سرانه پایین، صادرات سرانه نفتی کم و جمعیت زیاد هستند؛ در حالی که کشورهای گروه اول به توسعه حیاتی بلندمدت و پایدار می‌اندیشند، کشورهای با درآمد پایین تنها خواستار پوشش مشکلات مالی کوتاه‌مدت خود هستند (عبدلی و ناخدا، ۱۳۸۸).

در همین راستا، اقدام به تهدید در جهت بی‌ثباتی و تجدیدنظر در تصمیمات حکومت‌های خود می‌کنند. در مجموع می‌توان گفت از آنجا که کشورهای صبور از عامل تنزیل بالاتری

¹ Griffin & Vielhaber

نسبت به کشورهای کم‌صبر برخوردارند؛ لذا دارای موقعیت‌های چانه‌زنی مستحکم‌تری نیز می‌باشند. به عبارت دیگر، عامل تنزیل در گروه اول یا L بزرگ‌تر از گروه دوم S می‌باشد؛ $\delta_L > \delta_S$ (عبدلی و ماجد، ۱۳۹۱).

از سوی دیگر، نحوه تقسیم منافع حاصل از تشکیل ائتلاف، از طریق تعیین سهمیه تولید برای هر عضو صورت می‌پذیرد. لذا، بازیکنان برای دستیابی به سهمیه و سود فروش بیشتر با یکدیگر مذاکره و گاه مجادله می‌نمایند. از این‌رو، نحوه تقسیم منافع بین اعضای اوپک و استراتژی‌های مورد استفاده هر گروه در مواجهه با گروه مقابل و نیز نحوه عمل در مورد میزان تولیدات و تعاملات هر یک از اعضا با سایرین را می‌توان در چارچوب نظریه بازی‌ها مورد بررسی قرار داد.

۲-۲. پیشینه تحقیق

تاکنون مطالعات بسیاری مربوط به نوع شکل‌گیری اوپک و ظهور آن به عنوان نیروی غالب در بازارهای جهانی نفت با استفاده از رویکردهای متنوع مدل‌سازی کارتل صورت گرفته‌اند. برای مثال رویکرد نش-کورنو توسط پولاسکی^۱ (۱۹۹۲) برای مدل‌سازی اوپک به عنوان گروهی متحد پیشینه‌خواه و بدون لحاظ نزاع و اختلاف میان اعضا به کار گرفته شد.

اما پیندیک^۲ (۱۹۷۸) اوپک را به دو گروه کشورهای پس‌اندازکننده و مصرف‌کننده (کشورهای با نیاز شدید به نقدینگی) تقسیم کرده و توجه و تمرکز خود را بر قدرت چانه‌زنی این دو گروه معطوف ساخته است. در برخی مطالعات نیز مانند الحاجی و هوتنر^۳ (۲۰۰۰) با لحاظ عدم ناهمگونی بین اعضای اوپک، عربستان سعودی را به عنوان بنگاه غالب در کارتل معرفی کرده‌اند.

از سوی دیگر، برخی مطالعاتی که در چارچوب الگوی انحصاری به بررسی رفتار اوپک و طرف عرضه پرداخته‌اند، در قالب نظریه بازی‌ها صورت گرفته‌اند. دوتا^۴ (۱۹۹۹) با استفاده از رویکرد نظریه بازی به بررسی پویایی‌های درونی کشورهای اوپک طی زمان پرداخته است.

¹ Polasky

² Pindyck

³ Alhajji & Huettner

⁴ Dutta

گریفین و جیانگ^۱ (۱۹۹۷) در مطالعات خود نشان داده‌اند که با تشکیل کارتل و تبعیت از اصل همکاری در قالب بازی، منافع همه اعضای اوپک در مقایسه با وضعیت رقابتی افزایش می‌یابد. اما اعضاء همواره دارای این انگیزه‌اند که با فریب‌دادن دیگران و افزایش تولید مازاد بر سهمیه، منافع کوتاه‌مدت خود را افزایش دهند. اما ترس از رفتار تلافی‌جویانه سایر اعضاء در قبال فریب‌کاری آن‌ها و لذا کاهش منافع بلندمدت مانع از این اقدام می‌گردد.

آلت^۲ و همکاران (۱۹۹۸) گرچه مستقیماً به موضوع تقسیم منافع اوپک نپرداخته‌اند، اما به بحث درباره عکس‌العمل‌های استراتژیک بین تولیدکنندگان مهم نفتی پرداخته‌اند. آن‌ها استدلال می‌کنند که عربستان سعودی با داشتن هزینه‌های پایین تولید و ذخایر نفتی زیاد، به شرطی قادر به تحمل دوره‌های رکود قیمت نفت است که احتمال دهد از این بابت شهرتی نصیبش شده که به تبع آن کسب منفعت می‌کند. همچنین کشورهای نسبتاً ثروتمند اوپک مانند عربستان سعودی در زمینه مسئله همکاری اوپک به دنبال راه‌حلی بلندمدت و با ریسک کم هستند تا آن را جایگزین تنبیه اعضاء خاطی کنند.

عبدلی و ناخدا (۱۳۸۸) نیز در مطالعه خود با توجه به مدل چانه‌زنی و اجرای فیرون نشان داده‌اند که بازیکنان بی‌صبر اوپک پیامدهای بهتری نسبت به رقیبان صبورتر خود به دست می‌آورند. لذا چنین استدلال می‌کنند که دو نوع کشور در اوپک وجود دارد: گروهی با سرانه ذخایر بالا (کشورهای سازگار با آینده) و گروهی با سرانه ذخایر پائین (کشورهای سازگار با حال). نتایج بر اساس تفاوت نرخ تنزیل دو گروه کشورها حاکی از آن است که همکاری لزوماً در نتیجه افزایش مجازات اتفاق نمی‌افتد، بلکه گاهی اوقات افزایش پاداش‌ها روش کاراتری است که همکاری را موجب شود.

این نتایج در مطالعه عبدلی و ماجد (۱۳۹۱) نیز تأیید شده است. در این مطالعه، نظریه همکاری بین دو گروه ناهمگن اوپک در قالب بازی معمای زندانی از طریق برآورد مدل با آثار ثابت و استفاده از داده‌های تابلویی مورد سنجش واقع شده است. نتایج بیانگر آنند که در چانه‌زنی‌ها و مذاکرات، برخی اعضاء برای به‌سرعت به توافق رسیدن با باج‌دهی و کوتاه‌آمدن از مواضع سیاستی خود موجب تداوم عمر ائتلاف اوپک می‌گردند.

¹ Griffin & Xiong

² Alt et al.

ناجی میدانی و رحیمی (۱۳۹۵) نیز در راستای ارائه مکانیسم قیمت‌گذاری بهینه انرژی خصوصاً گاز بر اساس نظریه بازی‌های همکارانه و غیر همکارانه سناریوهای مختلفی تدوین و مدل‌سازی نموده‌اند.

با وجود این، در دهه اخیر با توجه به کارایی و کاربرد رویکردها و الگوریتم‌های تکاملی در حل مسائل کاربردی اقتصادی و تصمیم‌سازی، رویکرد به این روش رو به رشد است؛ به طوری که برخی محققان مانند دیوید^۱ (۱۹۹۹)، ریچمن^۲ (۲۰۰۱)، تویلز^۳ (۲۰۰۷) و جیتیس^۴ (۲۰۰۹) ضمن بررسی ادبیات موضوع EGT به ارائه کاربردهای اقتصادی آن پرداخته‌اند. پروان^۵ (۱۹۸۷) همچنین کاربردهای بازی تکاملی را برای هر تعداد از بازیکنان و مجموعه استراتژی‌های پیوسته به کار برد. ویگان^۶ و همکاران (۲۰۰۲) از EGT به عنوان ابزاری برای سنجش الگوریتم‌های تکاملی همکارانه استفاده کرده‌اند.

اما در زمینه به‌کارگیری نظریه بازی‌های تکاملی برای مدل‌سازی معمای زندانی، می‌توان مطالعه کاراندیکار^۷ و همکاران (۱۹۹۸) را برشمرد که با مدل‌سازی بازی‌های همکارانه و معمای زندانی تکاملی نتایجی مبنی بر رخداد همکاری با تناوب بیشتر نسبت به عدم همکاری در تکرار معمای زندانی را گزارش کردند.

پایه تحلیل تکاملی معمای زندانی تکراری را نیز می‌توان در مطالعه آکسلرود (۱۹۸۱) جستجو کرد. آکسلرود با استفاده از الگوریتم تکاملی به ارائه استراتژی‌های بهینه حل بازی پرداخته است. وی نخستین محقق بود که یافتن استراتژی‌های بهینه برای حل مسئله حداکثرسازی منافع را با استفاده از الگوریتم تکاملی (EA) مورد بررسی قرار داد. نتایج مطالعه وی حاکی از آن بود که استراتژی "همکاری" در بلندمدت در مقایسه با "عدم همکاری"، برنده بازی معمای زندانی خواهد بود.

روبنشتاین و آزبورن^۸ (۱۹۹۴) در مطالعه خود نتایج مسابقات آکسلرود را نقد کرده و

^۱ Dawid

^۲ Riechmann

^۳ Tuyls

^۴ Gintis

^۵ Brown

^۶ Wiegand

^۷ Karandikar

^۸ Rubenstein & Osborne

عوامل بسیاری را مانند طول کروموزوم، طول بازی و لحاظ استراتژی‌هایی به‌جز همکاری یا عدم همکاری محض را به عنوان استراتژی رقیب از جمله موانع بروز همه جانبه رفتار همکارانه در بازی عنوان نموده‌اند. این نقد و سایر مطالعات انتقادی از این دست منصفانه به قضاوت مطالعه آکسلرود پرداخته‌اند؛ اما مطالعه وی تاکنون به عنوان مناسب‌ترین رویکرد تکاملی شناخته شده است.

۳. مبانی نظری الگوریتم تکاملی برای مدل‌بندی اقتصادی

الگوریتم‌های تکاملی ابتدا برای استفاده در مسائل بهینه‌سازی طراحی شده و ادبیات گسترده‌ای نیز در زمینه تنظیم این نوع الگوریتم‌ها برای کارایی بهتر در مسائل بهینه‌سازی وجود دارد (بک و همکاران^۱، ۱۹۹۷؛ توسان و راس^۲، ۱۹۹۸)؛ اما استفاده از الگوریتم‌های تکاملی در حوزه شبیه‌سازی مسائل اقتصادی، بازنگری در دستورالعمل‌های پیشین را ناگزیر می‌سازد.

از جمله مهم‌ترین تمایزات بین مدل‌سازی اقتصادی و تکاملی می‌توان به نوع تنظیمات پارامترها در الگوریتم ژنتیک اشاره نمود؛ زیرا شبیه‌سازی‌های اقتصادی متعارف، اغلب تنظیمات پارامتری برای الگوریتم‌های ژنتیکی را مستقیماً از مقادیر پارامترهای مدل اقتصادی اقتباس می‌نمایند (دیوید، ۱۹۹۹)؛ در حالی که بر اساس مطالعه ریچمن (۲۰۰۱) این عمل کارایی GA و لذا یادگیری عوامل را محدود می‌نماید. در واقع در این نوع رویکرد، جمعیت تکاملی به عنوان جمعیتی از "عوامل" تلقی شده و لذا طی اجرای الگوریتم، این عامل‌ها هستند که تکامل می‌یابند. بنابراین، عامل‌های بهتر، طی فرایند تکاملی شناسایی شده و به نسل‌های بعدی به عنوان عاملان مولد منتقل می‌شوند. از این‌رو، با وجود جمعیت محدود، امکان همگرایی زودرس الگوریتم ژنتیک دور از انتظار نیست.

از سویی دیگر، در رویکرد اقتصادی، ارائه تعبیر اقتصادی مستقیم از یک پارامتر الگوریتم تکاملی (مانند اندازه جمعیت یا نرخ باز تولید) غالباً دشوار است. لذا به منظور حصول نتایج به لحاظ اعتباری قوی‌تر، پارامترهای مدل اقتصادی و پارامترهای الگوریتم تکاملی می‌بایست به طور جداگانه در نظر گرفته شوند. بنابراین، در رویکرد دوم، جمعیت تکاملی به عنوان جمعیتی

¹ Bäck et al.

² Tuson & Ross

از "استراتژی‌های" قابل انتخاب توسط عوامل شناخته می‌شوند (آلکمد، ۲۰۰۴). به عبارت دیگر، جمعیت کروموزوم‌ها همچون استخری از استراتژی‌هایی تلقی می‌شود که عوامل می‌توانند از بین آن‌ها انتخاب نمایند. از مزایای این امر می‌توان به انتخاب استراتژی‌های مشابه توسط چندین عامل و نیز بهبود کارایی الگوریتم تکاملی و یادگیری عوامل اشاره کرد.

جدول ۱. مقایسه الگوریتم‌های اقتصادی و تکاملی

رویکرد اقتصادی	رویکرد تکاملی
کروموزوم = عامل	کروموزوم = استراتژی
اندازه جمعیت توسط تعداد عوامل اقتصادی تعیین می‌شود	اندازه جمعیت از طریق فرایند یادگیری الگوریتم‌های تکاملی تعیین می‌شود

منبع: آلکمد (۲۰۰۴)

در الگوریتم تکاملی روند کار بدین صورت است که ابتدا جمعیتی از استراتژی‌های تصادفی اولیه تولید شده که در واقع جمعیت کروموزوم‌ها یا ژنوتایپ‌ها هستند. سپس رفتاری که هر عامل با به‌کارگیری یک استراتژی خاص بروز می‌دهد، روی کروموزوم کدگذاری می‌شود.

نحوه کدگذاری کروموزوم‌ها در اغلب مطالعات به‌صورت باینری (صفر و یک) می‌باشد. در روند تکامل، جمعیت کروموزوم‌های اولیه متعاقباً در نسل‌های بعدی از طریق انتخاب استراتژی بهتر با بازدهی تجمعی بیشتر و انتقال آن‌ها به نسل و جمعیت بعدی به عنوان مولد استراتژی‌های جدید دچار تغییر شده و بهبود می‌یابند. از این‌رو، بر مبنای اصل بقای اصلح، استراتژی‌هایی با برازندگی کم از جمعیت حذف شده و استراتژی‌های مولد و دارای بازدهی زیاد در جمعیت تکثیر می‌یابند (بک و همکاران، ۱۹۹۷). در مجموع، می‌توان گفت عمومیت الگوریتم‌های تکاملی در شبیه‌سازی‌های اقتصادی ناشی از توانمندی آنان در مدل‌سازی سیستم‌های بزرگی از عوامل با عقلانیت محدود و با رویکرد خرد به کلان^۱ است (میچل^۲، ۱۹۹۶).

^۱ From The Bottom up Approach

^۲ Mitchell

۳-۱. الگوریتم ژنتیک؛ بازی پویای تکاملی

الگوریتم ژنتیک، تکنیک محاسباتی تکاملی است که در سال ۱۹۸۹ توسط گلدبرگ^۱ با جزئیات معرفی گردید. در الگوریتم ژنتیک، راه حل ممکن یک مسئله به صورت رشته‌ای از بیت‌ها نشان داده می‌شود که طی فرایندی تصادفی مکرراً به یکدیگر تبدیل می‌شوند. این رشته‌های بیتی "افراد ژنتیکی" خوانده می‌شوند. به هر فرد ژنتیکی یک برازندگی تخصیص می‌یابد که بیانگر میزان کارایی آن فرد در حل مسئله پیش‌روست. با مقایسه برازندگی‌های حاصل توسط عوامل، نسل‌های جدید به‌طور مکرر با عملگرهای «انتخاب^۲»، «بازتولید^۳» و «جهش^۴» ایجاد می‌شوند. فرایندهای بازتولید و جهش، برخی از استراتژی‌های اقتصادی والدین را به‌منظور یافتن استراتژی‌های جدید (فرزندان) مورد استفاده قرار داده و آن‌ها را با استراتژی‌های فرزندان جایگزین می‌نمایند. لذا این امر موجب افزایش تنوع استراتژی‌ها درون جمعیت جاری می‌گردد. از سوی دیگر، عملگر ژنتیکی انتخاب، تعداد استراتژی‌های اقتصادی متفاوت را درون جمعیت کاهش می‌دهد. بدین‌صورت که ابتدا میزان بازدهی اقتصادی هر استراتژی را ارزیابی نموده و از این‌رو، اغلب به عنوان ایفاکننده نقش بازار یا آشکارکننده اطلاعات عمل می‌کند. سپس برخی از استراتژی‌ها را برای حضور در نسل بعدی برمی‌گزیند. شانس هر استراتژی i برای انتخاب از جمعیت جاری \bar{n} و تکثیر در دوره بعدی بستگی به برازندگی نسبی آن یعنی $P(i|\bar{n})$ ، به جمع بازدهی کلیه استراتژی‌ها در جمعیت یعنی $R(i|\bar{n})$ دارد؛ به‌طوری‌که بازدهی‌های نسبی بالاتر منجر به احتمال بازتولیدی بیشتر می‌شوند. بنابراین، تعداد استراتژی‌های متفاوت درون یک جمعیت با این فرایند مجدداً کاهش می‌یابد (ریچمن، ۲۰۰۱).

۳-۱-۱. الگوریتم ژنتیک اقتصادی به‌عنوان بازی پویا

الگوریتم‌های ژنتیک در واقع فرایندهای مدل‌سازی الگوریتم‌های یادگیری اجتماعی به واسطه تعاملات درون جمعیتی از عوامل اقتصادی هستند.

گفتنی است مشخصه مشترک هر دو موقعیت‌های نظری بازی و مدل‌های یادگیری GA وابستگی به وضعیت است. به‌علاوه، بازی در هر دور از الگوریتم ژنتیک تکرار شده و لذا به

¹ Goldberg

² Selection

³ Reproduction

⁴ Mutation

هر فرد شانس برای اصلاح و بهبود استراتژی خود داده می‌شود. بنابراین، می‌توان نتیجه‌گیری کرد که در واقع هر الگوریتم ژنتیک اقتصادی تک‌جمعیتی یک بازی پویاست؛ بدین صورت که در یک الگوریتم ژنتیک با جمعیتی متشکل از M فرد ژنتیکی که هر یک دارای رشته بیتی به طول L هستند، هر فرد ژنتیکی بر مبنای کدبندی باینری نمایانگر یکی از 2^L مقادیر متفاوت یا به عبارتی مجموعه استراتژی‌های در دسترس S است ($|S| = 2^L$). با توجه به اینکه در فرایند یادگیری الگوریتم ژنتیک، هیچ رقیب مستقیمی در برابر یک استراتژی منفرد وجود ندارد، لذا هر عامل اقتصادی در تلاش برای یافتن یک استراتژی $i \in S$ است که به‌خوبی هرچه تمام‌تر نسبت به تمامی جمعیت و محیط خود عمل کند.

۳-۱-۲. الگوریتم ژنتیک اقتصادی به‌عنوان بازی تکاملی

از آنجا که الگوریتم‌های ژنتیک به‌عنوان مدل‌هایی از یادگیری اقتصادی شناخته شده‌اند، آن‌ها را همچنین می‌توان به‌خوبی به‌عنوان فرایندهای تکاملی در نظر گرفت (دیوید، ۱۹۹۹).

در نگاه اول، در واقع این ساختار الگوریتم‌های ژنتیک و مدل‌های تکاملی است که مبین رابطه‌ای نزدیک بین GAها و نظریه اقتصادی تکاملی است؛ چرا که هر دو مواجه با ساختار مرکزی جمعیتی از عوامل اقتصادی متعامل در محیط‌های اقتصادی بوده و برای بهینه‌سازی رفتار فردی تلاش می‌کنند. به منظور ارائه شواهدی مبنی بر صحت این گزاره بر اساس تعریف فریدمن^۱ سه مشخصه برای یک بازی تکاملی ارائه می‌شود (فریدمن، ۱۹۹۸):

الف) در بازی‌های تکاملی، استراتژی‌های با بازدهی بالاتر طی زمان جایگزین استراتژی‌های با بازدهی کمتر می‌شوند. جایگزینی استراتژی‌ها در الگوریتم‌های ژنتیک فرایندی از تغییر در جمعیت ژنتیکی طی زمان است. این فرایند با استفاده از عملگرهای «انتخاب» و «بازتولید» و با تکثیر استراتژی‌های با بازدهی بالاتر به‌وقوع می‌پیوندد. لذا شرط اول فریدمن تأمین می‌شود.

ب) در طی فرایند جایگزینی استراتژی‌ها اینرسی وجود دارد. اینرسی به معنای آن است که تغییرات در رفتار عوامل به صورت بسیار ناگهانی اتفاق نمی‌افتند. در الگوریتم ژنتیک نیز عملگر جهش تنها منجر به تغییرات بسیار ناگهانی در بیت‌های تکی از رشته بیت کدگذاری

^۱ Friedman

شده مربوط به یک استراتژی اقتصادی با احتمال کوچک^۱ μ می‌گردد. از این‌رو، شرط دوم فریدمن نیز برقرار است.

ج) در بازی‌های تکاملی، بازیکنان به طور عمدی بر اقدامات آینده سایر بازیکنان تأثیرگذار نیستند. عوامل در یک مدل GA اقتصادی دانش بسیار محدودی دارند. هنگامی که یک عامل اقتصادی آخرین استراتژی اقتصادی‌اش را شکل می‌دهد درباره برنامه‌های سایر عوامل در جمعیت خود هیچ نمی‌داند. از این‌رو، همه آنچه که یک عامل اقتصادی در یک مدل GA می‌تواند انجام دهد، نهایت تلاش خود برای تطبیق با اقدامات گذشته رقبا یا به عبارت دیگر، انجام فرایند یادگیری است؛ لذا شرط سوم فریدمن نیز برقرار است. بنابراین، با توجه به موارد فوق می‌توان نتیجه گرفت که مدل‌های یادگیری GA اقتصادی را می‌توان به عنوان بازی‌های تکاملی تفسیر نمود. با این مزیت که الگوریتم‌های ژنتیک مفهوم روشنی از نفعی و حذف استراتژی‌های جهش و مهاجم از جمعیت را در مقایسه با مفهوم استراتژی‌های پایدار تکاملی در بازی‌های تکاملی ارائه می‌دهند (ویبول^۲، ۱۹۹۵).

۴. مدل‌سازی تکاملی IPD

برای حل بازی‌های تکراری خصوصاً IPD بر اساس مطالعه آکسلرود می‌توان از الگوریتم‌های تکاملی، مبتنی بر فرایند یادگیری عوامل طی هر دوره از بازی استفاده نمود. از این‌رو، بررسی و تحلیل تطابق رفتار عوامل اقتصادی یا بازیکنان بازی IPD در قالب الگوریتم‌های تکاملی در این بخش ارائه خواهد شد.

معمای زندانی یکی از مشهورترین بازی‌های استراتژیک است که به‌طور گسترده در علوم اقتصادی، سیاسی، یادگیری ماشینی و بیولوژی تکاملی مورد مطالعه قرار گرفته است. در این بازی هر دو بازیکن حرکت‌های خود را به‌طور همزمان و مستقل از انتخاب دیگری برمی‌گزینند.

در ماتریس بازدهی، زمانی که دو بازیکن با یکدیگر همکاری^۳ کرده به مقداری برابر R

^۱ احتمال جهش μ بطور نرمال عددی بین ۰/۰۰۱ تا ۰/۰۱ است.

^۲ Weibull

^۳ Cooperate (C)

پاداش داده می‌شوند. زمانی که تنها یک بازیکن همکاری را نقض کند^۱، وی بالاترین بازدهی ممکن (T) و رقیبش کم‌ترین بازدهی ممکن (S) را دریافت می‌کنند. همچنین در صورتی که هیچ یک از بازیکنان با دیگری همکاری ننمایند هر دو به میزان P جریمه خواهند شد. در ماتریس بازدهی $T > R > Q > S$ و نیز $R > (T + S)/2$ است. بدین معنا که وسوسه فریب‌کاری و نقض همکاری می‌بایست بازدهی بیشتری نسبت به همکاری داشته باشد.

جدول ۲. بازدهی‌های کلاسیک معمای زندانی

بازیکن ۲		استراتژی	
عدم همکاری (D)	همکاری (C)		
T=۵ S=۰	R=۳ R=۳	همکاری (C)	بازیکن ۱
P=۱ P=۱	S=۰ T=۵	عدم همکاری (D)	

منبع: تویلز، ۲۰۰۷

از آنجا که در این بازی، بازیکن اول (D,C) را به (C,C) و آن را به (D,D) و نهایتاً به (C,D) ترجیح داده و در مقابل بازیکن دوم (C,D) را به (C,C) و آن را به (D,D) و نهایتاً آن را به (D,C) ترجیح می‌دهد لذا تضاد موقعیتی بین بازیکنان ایجاد می‌شود (روبنشتاین و آزبورن، ۱۹۹۴). با حل بازی مشخص می‌شود که معمای زندانی یک تعادل نش اکید متقارن و منحصر به‌فرد عدم همکاری طرفین یا (D,D) دارد.

به عبارت دیگر، در هر تعادل کامل زیر بازی^۲، هر بازیکن استراتژی D را در هر بار بازی بدون در نظر گرفتن تاریخچه بازی اتخاذ می‌کند. لذا D تنها استراتژی پایدار تکاملی بازی است که در آن هیچ بازیکنی انگیزه‌ای برای تغییر استراتژی خود ندارد.

در حالی که اگر هر دو بازیکن بازی را برای تعداد دوره‌های نامتناهی تکرار کنند و بازدهی هر بازیکن برابر؛ مجموع تنزیل شده از بازدهی‌ها در هر نوبت از بازی باشد، آنگاه با انباشت بازدهی‌ها در هر دور ممکن است تعادلی بهتر از (D,D) وجود داشته باشد.

¹ Defect (D)

² Subgame Perfect Equilibrium

از این رو، رویکرد تکاملی IPD مبتنی بر الگوریتم ژنتیک، با بررسی تکامل استراتژی‌ها طی اجرای الگوریتم برای راندها و نیز نسل‌های مختلف این امر را ممکن می‌سازد. این رویکرد تکاملی مبتنی بر ارائه استراتژی‌ها به عنوان کروموزوم‌ها بوده که هر یک از آن‌ها حامل رفتاری کدگذاری شده از عوامل بازی می‌باشند. هدف هر اجرای الگوریتم ژنتیک، تکامل استراتژی‌های ممکن با بالاترین امتیاز یا بازدهی است. بدین منظور، در این مقاله در میان استراتژی‌های مطرح برای حل معمای زندانی تکراری، رقابتی ترتیب داده شده که در آن هر استراتژی در برابر دیگر استراتژی‌ها در بازی IPD قرار گرفته و برای کسب بازدهی بیشتر تلاش می‌کند.

بر اساس برازندگی حاصل توسط هر استراتژی در هر نسل، نسبتی از استراتژی‌های موفق توسط عملگر «انتخاب» GA برگزیده و با استفاده از عملگرهای «بازتولید» و «جهش» به نسل بعدی منتقل می‌شوند. بدین صورت، جمعیتی جدید از استراتژی‌ها برای اجراهای بعدی GA شکل می‌گیرند. در نهایت فرایند فوق در صورت لزوم به منظور سازگاری با مدل‌سازی IPD تکرار می‌گردد.

۴-۱. کدگذاری استراتژی‌های تکاملی

آکسلرود در مطالعه خود روشی ساده اما فوق‌العاده را برای کدگذاری استراتژی‌ها اتخاذ نمود که به عنوان روشی استاندارد از حل مسئله IPD در این مطالعه از آن استفاده می‌گردد. برای هر حرکت در بازی، چهار امکان وجود دارد که عبارتند از:

- هر دو بازیکن همکاری کنند (CC یا R برای پاداش)؛
- بازیکن دوم عدم همکاری در حالی که اولی همکاری کند (CD یا S برای فریفته شدن)؛
- بازیکن اول نقض همکاری و دومی همکاری کند (DC یا T برای فریب‌کاری)؛
- هر دو بازیکن همکاری نکنند (DD یا P برای تنبیه).

در واقع هر کروموزوم حاوی نتایج ممکن ۳ حرکت قبلی یک بازیکن است. لذا با برخورداری از اطلاعات سه حرکت قبلی یک بازیکن به منظور کدگذاری یک استراتژی خاص، رشته رفتاری مربوطه را در نظر گرفته و آن را به عنوان یک رشته سه حرفی کدگذاری

می‌نماییم. برای مثال، RRR ارائه‌گر نتیجه‌ای است که دو بازیکن در طی سه حرکت قبلی همکاری کنند و یا SSP بیان‌کننده نتیجه‌ای است که بازیکن اول دوبار با وجود عدم همکاری رقیب، همکاری کرده اما نهایتاً همکاری را نقض نموده است. سپس این رشته سه حرفی به منظور تولید عددی بین ۰ تا ۶۳ ($2 \times 2 \times 2 = 8$) مورد استفاده قرار گرفته و به‌عنوان عددی در مبنای ۴ تفسیر می‌گردد. یک راه ممکن جهت نمایش، تخصیص مقداری عددی به هر یک از کاراکترها به‌روشنی زیر است:

$$DD = P = 0 \quad DC = T = 1 \quad CD = S = 2 \quad CC = R = 3$$

از این‌رو، برای مثال در این روش، PPP به عدد صفر رمزگشایی شده و SSR به عدد $43 = (2 \times 4^2 + 2 \times 4^1 + 3 \times 4^0)$.

با دانش از ۳ حرکت قبلی برای حل این مسئله که بازیکن در حرکت جاری C یا D را بازی خواهد کرد، از الگوریتم ژنتیک (GA) استفاده می‌شود. بدین صورت که با وجود دو گزینه C یا D در هر ۶۴ موقعیت و مکان ممکن، یک استراتژی خاص می‌تواند به وسیله یک رشته باینری ۶۴ بیتی از C و D در الگوریتم ژنتیک تعریف شود که در آن، آامین C یا D متناظر با آامین رشته رفتاری دیکته شده توسط رشته سه حرفی از ۳ حرکت قبلی است. به عبارت دیگر، یک حرکت خاص، بستگی به ۳ حرکت قبلی خود دارد. هرچند ۳ حرکت اول در یک بازی به روش یاد شده تعریف نمی‌شوند. بلکه برای لحاظ این حرکت‌ها، ۶ بیت (با C و D اولیه تصادفی) به رشته ۶۴ بیتی فوق برای تعیین وضعیت یک استراتژی خاص با فرض در مورد رفتار آغازین بازی افزوده می‌گردد.

در مجموع می‌توان گفت هر رشته ۷۰ بیتی بیانگر یک استراتژی خاص است که ۶۴ بیت اول برای قاعده‌ها و ۶ بیت بعدی برای مکان‌ها استفاده می‌شوند. جدول (۳) یک رشته الگوریتم تکاملی (EA) نمونه را نشان می‌دهد. برای رشته نمونه در جدول یاد شده، کد سه حرفی PTP برای ۳ حرکت اولیه به‌دست می‌آید. این رشته به عدد ۴ رمزگشایی شده و به معنای آن است که بازیکن موردنظر می‌بایست $(4+1)$ یا پنجمین حرکت معین شده خود را در رشته ۶۴ بیتی الگوریتم ژنتیک بازی کند.^۱

^۱ از آنجا که مقادیر رمزگشایی از صفر آغاز می‌شوند، لذا با عدد یک جمع می‌شوند.

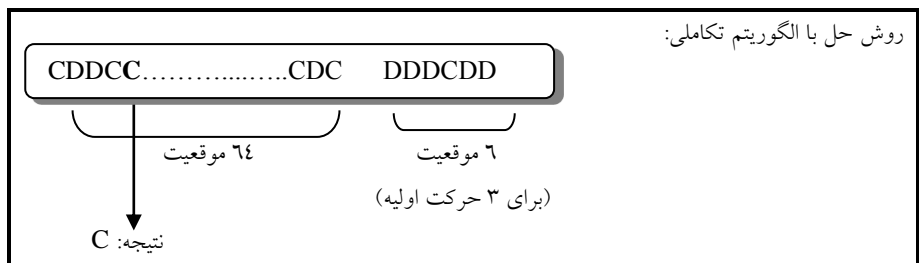
جدول ۳. کدگذاری یک استراتژی نمونه IPD

با وجود آگاهی از ۳ حرکت قبلی به شرح ذیل:			
کد گذاری	بازیکن دوم	بازیکن اول	
P	D	D	حرکت اول
T	C	D	حرکت دوم
P	D	D	حرکت سوم
نتیجه: بازیکن اول پنجمین موقعیت را انتخاب می کند $\Rightarrow 4 = (010)_4 = PTP$			

منبع: یافته‌های پژوهش

جدول (۴) روش حل رشته کدگذاری شده فوق را به روش تکاملی نشان می‌دهد. در این مورد بیت پنجم بیانگر C بوده، لذا بازیکن همکاری خواهد کرد.

جدول ۴. حل استراتژی کدگذاری شده به روش تکاملی



منبع: یافته‌های پژوهش

با استفاده از طرح کدگذاری فوق برای یافتن استراتژی‌های بهینه و نیز به‌کارگیری الگوریتم ژنتیک، می‌توان به استراتژی‌هایی با بیشترین بازدهی دست یافت.

۴-۲. شبیه‌سازی استراتژی‌های تکاملی

به منظور شبیه‌سازی استراتژی‌های مطرح در بازی معمای زندانی تکراری، ابتدا استراتژی‌ها کدگذاری شده و سپس در روند تکاملی به رقابت با یکدیگر می‌پردازند. طرح کدگذاری استفاده شده در این مقاله مشابه مطالعه آکسلرود است با این تفاوت که EA به کار گرفته شده

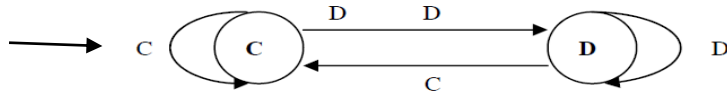
در مقاله حاضر یافتن دو نوع استراتژی بهینه هدف را مدنظر دارد. این دو هدف عبارتند از: (۱) حداکثرسازی امتیاز فردی و (۲) حداقل‌سازی امتیاز رقیب. باید گفت که امتیاز رقیب به معنای امتیاز تجمعی همه رقباست، زمانی که در مقابل یک استراتژی خاص بازی می‌نمایند. هر دو اجرای تک‌هدفه EA به منظور یافتن استراتژی‌های بهینه در این مطالعه صورت پذیرفته‌اند. شبیه‌سازی برای هر دو الگوریتم بر اساس ماتریس بازدهی جدول ۲ شامل مراحل زیر است:

در هر نسل، تعدادی معین از استراتژی‌ها تولید می‌شوند و هر استراتژی ملزم به بازی در برابر ۱۲ بازیکن (استراتژی) دیگر است. هر بازی از ۱۵۰ حرکت تشکیل شده است. به عبارت دیگر، هر فرد ژنتیکی معمای زندانی را برای ۱۵۰ بار در برابر ۱۲ رقیب در محیط خود بازی می‌کند.

برای محیط‌های ایستاتیک این امر به معنای بازی کردن در برابر تعداد از پیش تعیین‌شده استراتژی‌های شناخته شده است. در حالی که در محیط‌های تکاملی این به معنای ۱۵۰ بار بازی کردن با رقیب‌های به طور تصادفی انتخاب شده از زیر مجموعه رقباست. در پایان، استراتژی‌ها به تدریجی کاهشده، بر اساس امتیازات تجمعی آن‌ها دسته‌بندی شده و لذا نسل بعدی با استفاده از عملگرهای بازتولید و جهش شکل می‌گیرد.

در واقع، هدف هر اجرای الگوریتم ژنتیک تکامل استراتژی‌هایی با بالاترین بازدهی ممکن و انتقال استراتژی‌های کارا تر به نسل‌های بعدی است. استراتژی‌های تکاملی شبیه‌سازی شده در بازی IPD مطرح در این مطالعه عبارتند از:

۱. استراتژی همیشه همکاری (All C): در هر حرکت بازیکن همواره همکاری می‌کند.
۲. استراتژی همیشه نقض همکاری (All D): در هر حرکت بازیکن همواره نقض همکاری می‌کند.
۳. استراتژی مقابله به مثل (TFT): بازیکن در حرکت اول همکاری نموده و سپس آخرین حرکت رقیب را عیناً تکرار می‌کند.



شکل ۱. نمایش استراتژی TFT

۴. استراتژی TFT بدبین: مانند TFT بوده با این تفاوت که بازیکن در حرکت اول نقض همکاری می‌کند.
 ۵. استراتژی ماشه: بازیکن تا زمانی که سایر بازیکنان همکاری نمایند، همکاری نموده و در غیر این صورت همواره نقض همکاری می‌کند.
 ۶. استراتژی بازیکن تصادفی: بازیکن همکاری یا نقض همکاری را با احتمال یکسان به طور تصادفی انتخاب می‌کند.
 ۷. استراتژی بازی CD به طور تناوبی: بازیکن C و D را به صورت تناوبی بازی می‌کند.
 ۸. استراتژی بازی DDC به طور تناوبی: بازیکن D، D و C را به صورت تناوبی بازی می‌کند.
 ۹. استراتژی بازی CCD به طور تناوبی: بازیکن C، C و D را به صورت تناوبی بازی می‌کند.
 ۱۰. استراتژی TF2T یا عمل متقابل پس از دو بار نقض همکاری رقیب: مانند TFT بازیکن در حرکت اول همکاری نموده و پس از دو بار نقض همکاری رقیب، همکاری را نقض می‌کند.
 ۱۱. استراتژی اکثریت موافق: بازیکن در حرکت اول همکاری نموده و به بازی ادامه می‌دهد تا زمانی که تعداد دفعاتی که رقیب همکاری کرده بزرگ‌تر یا برابر با تعداد دفعات نقض همکاری او باشد، در غیر این صورت همکاری نمی‌کند.
 ۱۲. استراتژی اکثریت مخالف: در حرکت اول همکاری نکرده و به بازی ادامه می‌دهد تا زمانی که تعداد دفعاتی که رقیب نقض همکاری کرده بزرگ‌تر یا برابر با تعداد دفعات همکاری وی باشد، در غیر این صورت همکاری می‌کند.
- روشن است که در یک بازی خاص، با استفاده از مقادیر بازدهی نشان داده شده در

ماتریس بازدهی جدول ۲، یک بازیکن می‌تواند حداکثر امتیاز $5 \times 150 = 750$ را کسب کند. بدین‌صورت که وی همواره نقض همکاری و رقیب همواره همکاری نماید. همچنین حداقل امتیاز او چنانچه خود همیشه همکاری و رقیب همواره نقض همکاری نماید، برابر با صفر است. در حالی که هیچ‌یک از این دو حد آستانه معمولاً در عمل قابل دستیابی نیستند. بر اساس مطالعه دیوید (۱۹۹۹)، سنجش کاربردی‌تر در مورد یک استراتژی در بازی IPD عبارت است از محاسبه میزان نزدیکی آن به امتیاز معیار^۱.

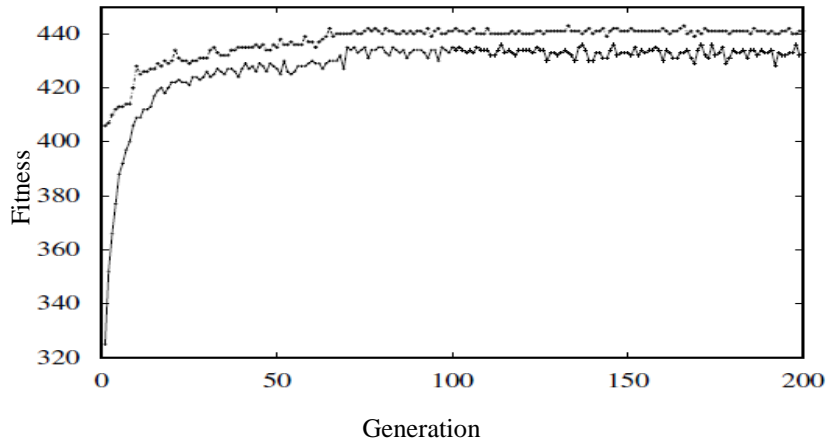
بنا بر تعریف، امتیاز معیار، امتیازی است که در آن هر دو بازیکن همواره همکاری نمایند. با توجه به جدول (۲) مسئله IPD، امتیاز معیار برابر $3 \times 150 = 450$ در این مورد است. برای مثال اگر امتیاز یک بازیکن به طور میانگین در برابر همه بازیکنانی که با آن‌ها بازی کرده ۴۰۰ باشد، آنگاه می‌توان گفت که او ۸۹ درصد از امتیاز معیار را کسب نموده است. این روش برای برشمردن امتیازات یک بازیکن بسیار مناسب بوده است؛ زیرا مستقل از مقادیر موجود در ماتریس بازدهی‌های استفاده شده و نیز تعداد بازیکنان رقیب می‌باشد. از این‌رو، در تمامی نتایج ارائه شده در این مطالعه امتیازات به صورت درصدی از امتیاز معیار در نظر گرفته شده‌اند.

۵. نتایج شبیه‌سازی بازی IPD با استفاده از الگوریتم تکاملی

دو اجرای مستقل الگوریتم تکاملی در این مطالعه به منظور حداکثرسازی امتیاز فردی بازیکن و نیز حداقل‌سازی امتیاز رقیب به‌کار رفته است. از آنجا که راه‌حل‌ها با استفاده از رشته بیت ارائه می‌شوند؛ لذا از عملگرهای «انتخاب باینری»، عملگر «تقاطع تک نقطه‌ای^۲» و عملگر «جهش بیتی» استفاده کرده‌ایم. بدین صورت که در تمامی شبیه‌سازی‌ها احتمال تقاطع $0/9$ ، و احتمال جهش $1/70$ می‌باشند. بنابراین به طور متوسط تنها یک بیت در یک رشته ۷۰ بیتی در واحد زمان جهش می‌یابد. برای هر اجرا اندازه جمعیت متناظر با ۴۰ و تعداد نسل‌ها ۲۰۰۰۰ نسل تنظیم گردید. نمودار امتیازات برای ۲۰۰ نسل اول، با اجرای EA در نمودارهای (۱ و ۲) نشان داده شده‌اند. با این تفاوت که در نمودار (۱) امتیاز فردی حداکثر شده و در نمودار (۲) امتیاز رقیب حداقل می‌گردد.

¹ Benchmark score

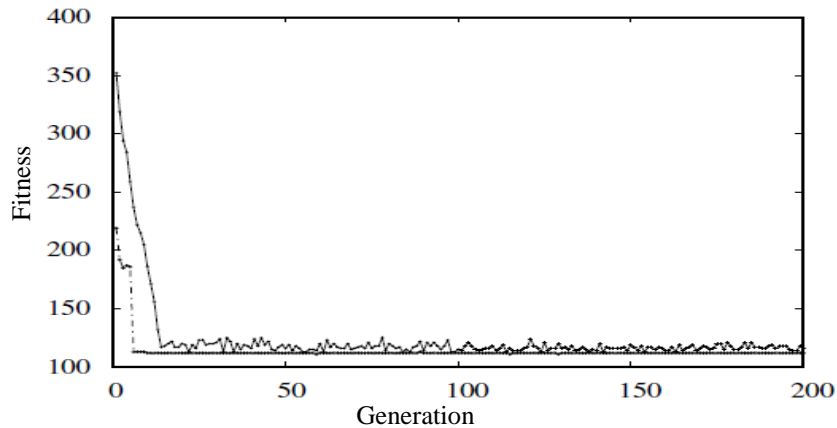
² Single-Point Crossover Operator



نمودار ۱. برازندگی متوسط (نمودار زیرین) و حداکثر برازندگی فردی (نمودار فوقانی)

برای حداکثرسازی امتیاز فردی با استفاده از EA، سنجش برازندگی نمونه بر اساس امتیاز فردی است. از این رو، امتیازات مربوط به برازندگی می‌بایست حداکثر گردد. همان‌طور که از نمودار (۱) مشخص است؛ در مورد اول برازندگی متوسط به‌طور یکنواخت طی نسل‌ها افزایش یافته و در حدود نسل ۲۰۰، امتیاز فردی با برازندگی ۴۴۱ حداکثر شده و بدین ترتیب ۹۸ درصد امتیاز معیار حاصل می‌گردد. زمانی که EA برای نسل‌های بیشتر اجرا می‌شود، حداکثر برازندگی به عدد ۴۴۶ (۹۹/۱٪ امتیاز معیار) همگرا شده و دیگر بیش از آن افزایش نمی‌یابد. این روند در مطالعات بهینه‌سازی با استفاده از GAها معمول است. بدین صورت که روند فزاینده برازندگی در ابتدا سریع و در ادامه نسبتاً کند اتفاق می‌افتد. زیرا یافتن راه‌حل‌های بهتر نزدیک به راه‌حل بهینه به‌مرور در مسائل بهینه‌سازی دشوارتر می‌شود.

زمانی که امتیاز رقیب با الگوریتم تکاملی حداقل می‌گردد، کمترین برازندگی در ۱۱۲ یعنی ۲۴/۹ درصد از امتیاز معیار به‌دست می‌آید. این نتیجه مبنی بر حداقل کارایی و مشابه با اتخاذ استراتژی همواره عدم همکاری است.



نمودار ۲. برازندگی متوسط (نمودار فوقانی) و حداقل برازندگی رقیب (نمودار زیرین)

برای مقایسه امتیازات استراتژی‌های مطرح در IPD نفتی، جدول‌های ۵ و ۶ نتایج را برای ۲۰ اجرای GA از هر دو رقابت نشان می‌دهند. در رقابت نخست ۱۲ استراتژی شرکت نموده و "TFT" استراتژی برنده با امتیاز متوسط ۳۸۵ (۸۵٪ امتیاز معیار) شناخته می‌شود. دلیل این امر توانمندی استراتژی TFT در پاداش دادن به رقیب همکاری‌کننده و نیز تنبیه وی به دلیل عدم همکاری‌اش است. در واقع، تکرار بازی، وجود استراتژی TFT را همانند تعاملات در دنیای واقعی، ممکن می‌سازد؛ زیرا بازیکنان قادر خواهند بود تا بر اساس رفتارهای پیشین یکدیگر تصمیم به پاداش یا تنبیه رقیب نمایند.

زمانی که امتیاز رقیب به‌تنهایی با EA حداقل می‌گردد، باز هم استراتژی TFT به عنوان استراتژی با بیشترین بازدهی در این رقابت معرفی می‌شود؛ در حالی که استراتژی همیشه عدم همکاری (All D) به عنوان ناکارترین استراتژی با پایین‌ترین میزان بازدهی شناخته می‌شود. سایر استراتژی‌ها که به‌خوبی TFT حائز امتیازات بالا هستند، عبارتند از: استراتژی اکثریت موافق، استراتژی ماشه و نیز استراتژی TF2T که گونه‌ای مشتق شده از استراتژی TFT است. همچنین، استراتژی همواره عدم‌همکاری در رقابت‌ها با کمترین بازدهی به عنوان ناکارترین استراتژی شناخته شده است.

جدول ۵. امتیازات استراتژی‌های IPD برای حداکثرسازی امتیاز فردی

رتبه	استراتژی بازیکن	امتیاز متوسط
۱	TFT	۳۹۰
۲	اکثریت موافق	۳۷۵
۳	TF2T	۳۶۷
۴	استراتژی ماشه	۳۶۳
۵	همیشه همکاری	۳۶۰
۶	بازی تناوبی CCD	۳۵۵
۷	بازی تناوبی CD	۳۳۵
۸	اکثریت مخالف	۳۳۰
۹	بازیکن تصادفی	۳۱۸
۱۰	TFT بدبین	۳۱۰
۱۱	همیشه عدم همکاری	۲۹۶
۱۲	بازی تناوبی DDC	۲۹۵

منبع: یافته‌های پژوهش^۱

جدول ۶. امتیازات استراتژی‌های IPD برای حداقل سازی امتیاز رقیب

رتبه	استراتژی بازیکن	امتیاز متوسط
۱	TFT	۳۸۵
۲	اکثریت موافق	۳۷۵
۳	استراتژی ماشه	۳۷۳
۴	TF2T	۳۷۲
۵	همیشه همکاری	۳۶۴
۶	بازی تناوبی CCD	۳۵۵
۷	بازی تناوبی CD	۳۴۸
۸	بازیکن تصادفی	۳۳۰

^۱ با استفاده از نرم افزار Matlab

رتبه	استراتژی بازیکن	امتیاز متوسط
۹	اکثریت مخالف	۳۱۵
۱۰	بازی تناوبی DDC	۳۱۰
۱۱	TFT بدبین	۳۰۸
۱۲	همیشه عدم همکاری	۳۰۶

منبع: یافته‌های پژوهش

۶. جمع‌بندی و نتیجه‌گیری

در محیط‌های پیچیده، افراد به طور کامل قادر به تحلیل موقعیت‌ها و اتخاذ استراتژی‌های بهینه خود نیستند. اما چنین انتظار می‌رود که بتوانند استراتژی خود را طی زمان بر اساس اینکه کدام یک موثر و کدام ناکاراست، تطبیق دهند. قیاسی مناسب برای فرایند انطباق، تکامل بیولوژیکی است. در تکامل، استراتژی‌های نسبتاً موثر گسترده‌تر شده و استراتژی‌های کمتر مؤثر به تدریج از جمعیت حذف می‌گردند.

شایان ذکر است که نگرش مقاله حاضر به بررسی نحوه تکامل و ظهور همکاری در موقعیت‌های رقابتی اختصاص داشته و هدف آن ارائه مدلی جدید از جستجوی استراتژی‌های بهینه در بازی معمای زندانی تکراری است. لذا بر مبنای توانمندی فرایند تکامل و الگوریتم‌های تکاملی در تولید استراتژی‌هایی با کارایی بهتر در بازی‌های تکراری، در این مطالعه به بررسی استراتژی‌های بهینه در بازی معمای زندانی تکراری (IPD) با استفاده از الگوریتم ژنتیک پرداخته شده است. بدین منظور با شرکت دو کشور ایران و عربستان به عنوان نمایندگان دو گروه ناهمگن اوپک و لحاظ ۱۲ نوع استراتژی مطرح در IPD، بازی تکاملی با دو رویکرد؛ حداکثرسازی امتیازات فردی و نیز حداقل‌سازی امتیاز رقیب شبیه‌سازی و اجرا گردید.

نتایج دو رقابت بین ۱۲ نوع استراتژی با رویکرد تکاملی حاکی از آن است که استراتژی "عمل متقابل" یا "TFT" به عنوان استراتژی بهینه حائز بالاترین بازدهی در حداکثرسازی امتیازات فردی بازیکن و نیز حداقل‌سازی امتیاز رقیب وی شناخته شده است. این استراتژی در رقابت‌ها، به دلیل توانمندی منحصر به فرد خود در پاداش دادن به رقیب همکاری‌کننده و نیز

تنبیه وی به دلیل عدم همکاری اش امتیاز بالایی کسب می‌کند. دستیابی به استراتژی بهینه در رقابت دو گروه ناهمگن اوپک دال بر همکاری اولیه و سپس مقابله به مثل رفتاری است. اتخاذ چنین استراتژی ضمن حفظ موقعیت بازیکن، با ایجاد تهدید قابل قبول و معتبر مانع رخداد تنازع و عدم همکاری بین اعضاء خواهد شد.

سایر استراتژی‌ها که به خوبی TFT حائز امتیازات بالا هستند عبارتند از: استراتژی اکثریت موافق، استراتژی ماشه و نیز استراتژی TF2T که گونه‌ای مشتق شده از استراتژی TFT است. همچنین، استراتژی همواره عدم همکاری در رقابت‌ها با کمترین بازدهی به عنوان ناکارترین استراتژی شناخته شده است.

منابع

- جمشیدی رودباری، مستانه (۱۳۸۷). بررسی علل تطابق نیافتن مدل‌های اقتصادی رفتار اوپک در بلندمدت از دیدگاه تحولات بازار نفت و ویژگی‌های این سازمان. *فصلنامه پژوهش‌ها و سیاست‌های اقتصادی*، ۴۷: ۶۳-۲۵.
- سامتی، مرتضی، فتح‌آبادی، مهدی، کسرای، کامران (۱۳۹۰). تعادل استراتژی مختلط نش و بازیکنان فوتبال: مطالعه موردی ضربات پنالتی. *فصلنامه مدل‌سازی اقتصادی*، ۱۵(۵): ۶۶-۴۷.
- عبدلی، قهرمان، ماجد، وحید (۱۳۹۱). بررسی رفتار اوپک در قالب یک بازی همکارانه. *تحقیقات مدل‌سازی اقتصادی*، ۲(۷): ۵۰-۲۷.
- عبدلی، قهرمان، ناخدا، محمد جواد (۱۳۸۸). کاربرد نظریه فیرون در بررسی پایداری اوپک: با رویکرد نظریه بازی‌های تکراری. *فصلنامه مطالعات اقتصاد انرژی*، ۶(۲۰): ۵۶-۳۳.
- ناجی میدانی، علی اکبر، رحیمی، غلامعلی (۱۳۹۵). مدل قیمت‌گذاری صادرات گاز طبیعی از طریق خط لوله بر اساس نظریه بازی‌ها. *فصلنامه مدل‌سازی اقتصادی*، ۲(۳۴): ۴۹-۲۹.
- Alhajji, A. F., & Huettner, D. (2000). OPEC and world crude oil markets from 1973 to 1994: Cartel, Oligopoly, or Competitive? *The Energy Journal*, 21(3): 31-60.
- Alkemade, F. (2004). Evolutionary agent-based economics. Eindhoven: Technische Universiteit Eindhoven.

- Alt, J. E., Calvert, L., & Humes, B. D. (1988). Reputation and Hegemonic Stability: A Game-Theoretic Analysis. *American Political Science Review*, 82(2): 445-66.
- Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. *Science*, 211(4489): 1390-96.
- Back, T., Fogel, D., & Michalewicz, Z. (1997). Handbook of evolutionary computation. Oxford University Press.
- Brown, J. S. (1987). A theory for the evolutionary game. *Theoretical Population Biology*, 31(1): 140-166.
- Dawid, H. (1999). Adaptive learning by genetic algorithms: analytical results and applications to economic models. Springer Verlag, Berlin. 2nd Edition.
- Duboz, R., Versmisse, D., Travers M., Ramat, E., & Shin, Y. J. (2010). Application of an evolutionary algorithm to the inverse parameter estimation of an individual-based model. *Ecological Modelling*, 221(5): 840-849.
- Dutta, P. K. (1999). *Strategies and Games: Theory and Practice*. The MIT Press.
- Friedman, D. (1998). On economic applications of evolutionary game theory. *Journal of Evolutionary Economics*, 8(1): 15-43.
- Gintis, H. (2009). Game theory evolving: a problem-centered introduction to modeling strategic interaction. Princeton, NJ: Princeton University Press.
- Goldberg, D. E. (1989). Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Reading, MA.
- Griffin, J., & Vielhaber, L. (1994). OPEC production: the missing link. *The Energy Journal*, 15:32- 115.
- Griffin, J. M., & Xiong, W. (1997). The Incentive to Cheat: An Empirical Analysis of OPEC. *Journal of Law and Economics*, 40(2): 289-316.
- Karandikar, R. Mookherjee, D., Ray, D., & Vega- Redondo, F. (1998). Evolving aspirations and cooperation. *Journal of Economic Theory*, 80(2): 292-331.
- Marks, R. E. (1992). Breeding hybrid strategies: Optimal behaviour for oligopolists. *Journal of Evolutionary Economics*, 2, 17- 38.
- Mitchell, M. (1996). An introduction to genetic algorithms. London MIT Press, Cambridge, MA.
- Pindyck, R. S. (1978). Gains to producers from the cartelization of exhaustible resources. *The Review of Economics and Statistics*, 60(2): 238-251.
- Polasky, S. (1992). Do oil producers act as 'Oil'igopolists? *Journal of Environmental Economics and Management*, 23(3): 216-247.

- Riechmann, T. (2001). Genetic algorithm learning and evolutionary games. *Journal of Economic Dynamics and Control*, 25(6): 1019-1037.
- Rubenstein, M., & Osborne, M. (1994). A course in game theory. Cambridge, MA: MIT Press.
- Smith, J. M. (1982). Evolution and the theory of games. Cambridge, UK: Cambridge University Press.
- Tuson, P., & Ross, P. (1998). Adapting operator settings in genetic algorithms. *Evolutionary computation*, 6(2): 161-184.
- Tuyls, K., & Parsons, S. (2007). What evolutionary game theory tells us about multiagent learning. *Artificial Intelligence*, 171(7): 406-416.
- Von Neumann, J., & Morgenstern, O. (1944). Theory of games and economic behavior. Princeton University Press.
- Weibull, J. W. (1995). Evolutionary game theory. London: MIT Press, Cambridge, MA.
- Wiegand, R. P., Liles, C. L., & De Jong, K. A. (2002). Analyzing Cooperative Coevolution with Evolutionary Game Theory. In *Proceedings of Congress on Evolutionary Computation (CEC-02)*, edited by D. Fogel. IEEE Press.

