دوفصلنامه بازی جنگ، سال سوم، شماره ۷، اسفند ۱۳۹۹

تاریخ دریافت: ۱۳۹۹/۰۲/۲۵ تاریخ پذیرش: ۱۳۹۹/۰۸/۱۵

از صفحه ۹۷ تا صفحه ۱۳۱

بررسی انواع روش‌های بازی در نظریه بازی‌ها

نوع مقاله: ترویجی

پژمان غلام نژاد*[1]

چکیده

در دهه اخیر، به کارگیری نظریه بازی‌ها در علوم مختلف، رشد چشمگیری داشته است. با توجه به رشدِ روز افزونِ فناوری اطلاعات و هوش مصنوعی در میدان‌های نبرد، منافع استراتژیکِ هر یک از طرفین، بستگی زیادی به عملکردِ هوشمندانه طرف مقابل دارد. به دلیلِ تغییرِ شکلِ ماهیتِ نبردها که عمدتا مبتنی بر فناوری اطلاعات و هوش مصنوعی هستند، جایگزینیِ روش‌های نوین در این حوزه، می‌تواند نتایجِ اثربخشی را در بر داشته باشد. در واقع، هر جا که منابع محدود، گزینه‌های مختلفِ تصمیم گیری، دستاوردهایِ متفاوت در اثر انتخاب‌های متفاوت و امکان همکاری یا رقابت بین بازیگران وجود داشته باشد، می‌توان از نظریه بازی ها به منظورِ تحلیلِ بهترِ شرایطِ موجود استفاده کرد. هدف از این تحقیق، معرفی و تعریف نظریه بازی‌ها، روش‌های نمایش و انواع آن می‌باشد تا بتوان از نقشِ نظریه بازی‌ها در میدانِ نبرد، استفاده بهینه نمود. روش تحقیق، بر اساس ماهیت و نحوه گردآوری داده‌های آن، توصیفی (موردی و زمینه‌ای) است، زیرا قرار است در یک مورد خاص (انواع بازی‌ها در نظریه بازی‌ها) عمیقا پژوهش به عمل آید. مهمترین نتایج حاصل از این تحقیق، شناخت و به کارگیری انوعِ روش‌های مبتنی بر نظریه بازی‌ها می‌باشد که می‌تواند منجر به افزایشِ بازدارندگیِ دفاعی در میدان‌های نبرد گردد.

واژگان کلیدی:

نظریه بازی‌ها، هوش مصنوعی، انواع بازی‌ها، فضای بازی.

[1] دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه علوم و فنون هوایی شهید ستاری، تهران، ایران.

* نویسنده مسئول: pezhman.gholamnezhad@gmail.com

مقدمه

امروزه پیشـــرفتِ انقلابی در فناوری اطلاعات و ارتباطات، تغییراتِ زیادی را در ماهیتِ نبردها ایجاد کرده اسـت و اسـتفاده از این فناوری به عنوان ابزاری اسـاسـی در برتری و قدرت نظامی محسـوب می شـود. فضـای سـایبر به عنوان مولود فناوری اطلاعات و ارتباطات، میدان جنگ سـنتی را به جنگ‌های نوین تبدیل نموده و پس از زمین، دریا، هوا و فضـا به عنوان پنجمین صحنه نبرد شناخته می شود. به خاطر وابستگی اکثر زیر ساخت‌های حیاتی کشورها به فضای سایبر، این فضا در کنار مزایای بی‌شمار برای جوامع بشری، تهدیدات نوینی نیز به همراه آورده است.

در فضای سایبر به واسطه از بین رفتن مرزهای فیزیکی، حملات سایبری، متعددی از جمله حملات مهندسی اجتماعی، ویروس‌های بروز شده، اسب‌های تراوا، کرم‌ها، حمله محروم‌سازی از سرویس[1]، حمله محروم سازی از سرویسِ توزیع شده[2]، بات‌نت[3] و ... به وجود آمده ا ست. این حملات، زیر سـاخت‌های حیاتی کشـورها را تهدید می‌کند و مناقشـات بین کشـورها از فضـای واقعی به فضای سایبر کشیده شده است. پیچیدگی حملات سایبری، دفاع در برابر این حملات را نیز مشکل ساخته است.

از آن‌جا که سامانه‌های دفاعی نوین، وابستگی شدیدی به فضای سایبر دارند، هرگونه نفوذ، خرابکاری و افشای اطلاعات سازمان‌ها، هزینه‌های زیادی را در پی خواهد داشت بنابراین، امنیت شبکه‌ها و سامانه‌های سخت افزاری و نرم افزاری در حوزه دفاعی و نظامی حائز اهمیت است. با توجه به افزایشِ سرعتِ پردازشِ داده‌ها در دفاع سایبری، نمی‌توان انتظار داشت که بدون به کارگیری فناوری اطلاعات و نظریه بازی‌ها و استفاده‌ی تنها از نیروی انسانی، به یک دفاعِ موثر دست یافت. همچنین دفاع در برابرِ حملاتِ در حالِ تحولِ پویا مبتنی بر ابزارهای هوش مصنوعی، با استفاده از الگوریتم‌های نرم‌افزاری ثابت و معمولی، دشوار و گاها غیر ممکن است.

در میدان نبردِ واقعی، منافعِ اسـتراتژیک مهاجم (مدافع)، بسـتگی زیادی به عملکردِ مدافع (مهاجم) دارد. بنابراین اثربخشیِ سازوکار دفاعی، متکی به رفتارهای استراتژیکِ مدافع و مهاجم اسـت و به کارگیریِ یک ابزارِ موثر، مبتنی بر فناوری اطلاعات، در این امر، می‌تواند در افزایشِ راندمانِ این اثربخشی، چشم‌گیر باشد. شناخت و به کارگیریِ ابزارها و روش‌های مبتنی بر هوش

[1] Denial of Service Attack (DoS)
[2] Distributed Denial of Service Attack (DDoS)
[3] Botnet

مصـــنوعی و نظریه بازی‌ها در این میدانِ نبرد، می‌تواند منجر به افزایش چشـــم‌گیرِ بازدارندگی دفاعی، در میدان‌های نبرد شود.

نظریه بازی، زیرمجموعه‌ای از علم ریاضیات است که تلاش می‌کند با استفاده از طراحی و تحلیل سناریو، رفتارها و نتایج تصمیم‌گیری موجوداتی را که حق انتخاب دارند، در تعامل با یکدیگر پیش‌بینی کند. در واقع، هر جا که منابع محدود، گزینه‌های مختلف تصمیم گیری، دستاوردهای متفاوت در اثر انتخاب‌های متفاوت و امکان همکاری یا رقابت بین بازیگران وجود داشته باشد، می‌توان از نظریه بازی‌ها برای درک و تحلیل بهتر شرایط موجود استفاده کرد.

اهمیت و ضرورت تحقیق

شـــناخت انواع بازی‌ها، می تواند با تجزیه و تحلیل تاکتیکی، و بررســـیِ موقعیت‌های تصمیم‌گیریِ استراتژیکِ مدافع و یا تجزیه و تحلیل انگیزه‌های مهاجمان، برتری تصمیم سازی در صحنه نبرد را به همراه داشته باشد.

تولید محتوای علمیِ، در حوزه نظریه بازی، برای پیش بینی، بســـیاری از مخاطرات امنیت سایبری در سامانه‌های دفاع سایبری از اهمیت بالایی برخوردار است.

با توجه به غلبه‌ی رویکردِ نظریه بازی بر راه حل‌های سنتیِ سایبری (ریاضیاتِ اثبات شده، دفاعِ قابلِ اعتماد، اقدامِ به موقع و راهِ حل‌های توزیع شده)، بیانِ نقشِ تئوری بازی در دست‌یابی به یک استراتژیِ تعادل برای زنده ماندن از حملاتِ غیر منتظره، می‌تواند در تصمیم‌گیری‌های آتی، بسیار اثربخش باشد.

فقدان دانش کافی و تحقیقات علمی در زمینه رویکردهای سامانه‌های دفاعی سایبری که ترکیبی از سامانه‌های فیزیکی سایبری است؛ می تواند منجر به چالش‌ها و مشکلات امنیتی شود.

هدف تحقیق

هدف اصلی این تحقیق، معرفی نظریه بازی‌ها، نمایش آن‌ها و انواع بازی‌ها می‌باشد.

پیشینه تحقیق

در سال ۲۰۰۴، جووِ کامبرا[1] و همکارانش استفاده از بازی‌های آنلاین چند نفره را برای ارزیابی سامانه‌های فرماندهی و کنترل پیشنهاد نمودند[۱]. در سال ۲۰۰۷، جُزِف وای. هالپِرن[2] و همکارانش به بررسی روابط بین علوم کامپیوتر و نظریه بازی‌ها پرداختند[۲]. آن‌ها نقش پیچیدگی محاسباتیِ نظریه‌بازی‌ها را بررسی نمودند و یک مساله نظریه بازی‌ها را که از علوم کامپیوتر

1 Juve Kambra
2 Joseph Y. Halpern

الگوبرداری شده است را معرفی نمودند. در سال ۲۰۱۰، سَنکارداس رُی[1] و همکارانش به بررسی کاربرد روش‌های نظریه‌بازی‌ها در امنیت شبکه پرداختند[۳] و بر اساس رویکردهای نظریه‌بازی، راه حل‌های پیشنهادی را در امنیت شبکه بیان نمودند. در سال ۲۰۱۲، های یان شی[2] و همکارانش به بررسی کاربردهای نظریه بازی در شبکه‌های حسگر بی‌سیم پرداختند[۴]. در سال ۲۰۱۶، یووان وانگ[3] و همکارانش به بررسی روش‌های نظریه بازی در کاربردهای امنیت سایبری[4] براساس طبقه‌بندی های امنیتی پرداختند[۵]. در سال ۲۰۱۷، کونگ دی دو[5] و همکارانش، رویکردهای نظریه بازی را برای امنیت سایبری و مسایل مربوط به حریم خصوصی را بررسی نمودند[۶]. در سال ۲۰۱۷، گوانچی یانگ[6] و همکارانش، به منظور بهینه‌سازی، الگوریتم‌های تکاملی مبتنی بر نظریه بازی را بررسی نمودند[۷]. در سال ۲۰۱۹، اِم دی عَرَفات حبیب[7] و همکارانش یک بررسی مقایسه‌ای برای مسیریابی مبتنی بر تئوری بازی برای شبکه‌های حسگر بی‌سیم انجام دادند[۸]. در سال ۲۰۱۹، زیاو لیو[8] و همکارانش، به بررسی کاربردهای نظریه بازی در بلاک چین[9] پرداختند[۹]. در سال ۲۰۱۹، سارا ریاحی و همکارانش، به نقش نظریه باری، برای به اشتراک گذاری منابع در سیستم‌های توزیع شده بزرگ پرداختند[۱۰]. در سال ۲۰۲۰، مارگاریتا زرگریان[10]، به بررسی نظریه بازی‌ها در الگوریتم‌های توزیع شده به منظور تولید بازی-های مبتنی بر موبایل پرداخت[۱۱].

نوع و روش تحقیق

این تحقیق از نظر هدف توسعه‌ای است، زیرا به دنبال یافتن روشی علمی مناسب برای حل یک مسئله است. از نظر ماهیت داده‌ها آمیخته (کمی و کیفی) و از نظر روش گردآوردی داده ها نیز توصیفی است. روش تحقیق در این پژوهش بر اساس ماهیت و نحوه گردآوری داده‌های آن، توصیفی (موردی و زمینه‌ای) است. با رویکرد آمیخته (کمی و کیفی) است، زیرا قرار است در یک مورد خاص (انواع بازی‌ها در نظریه بازی‌ها) عمیقا پژوهش به عمل آید.

---

[1] Sankardas Roy
[2] Hai-Yan Shi
[3] Yuan Wang
[4] Cyber Defense
[5] Cuong T Do
[6] Guanci Yang
[7] Md Arafat Habib
[8] Ziyao Lio
[9] Block Chain
[10] Margarita Zargaryan

نظریه بازی‌ها

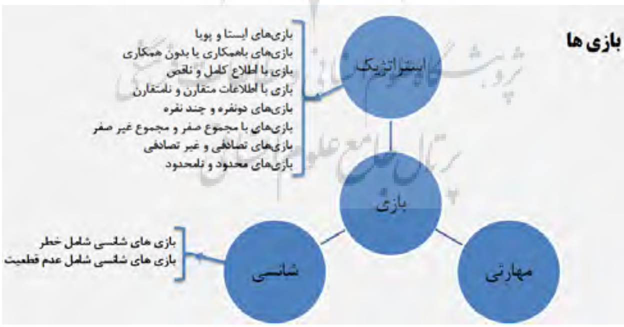
یک بازی شامل مجموعه‌ای از بازیکنان، مجموعه‌ای از حرکت‌ها یا راه بردها و نتیجه‌ی مشخصی برای هر ترکیب از راه بردها می‌باشد. پیروزی در هر بازی تنها تابع شانس نیست، بلکه اصول و قوانینِ ویژه‌ی خود را دارد و هر بازیکن در طی بازی سعی می‌کند با به‌کارگیری آن اصول، خود را به بُرد نزدیک کند. بنابراین هر بازیکن هنگام تصمیم‌گیری برای انجام حرکت بهینه، بایستی تمامی واکنش‌های بازیکان دیگر را نسبت به حرکتِ خود، در نظر بگیرد، در حالی که حرکات سایر بازیکنان را با قطعیت نمی‌داند. اما درباره‌ی حرکت خود بایستی تصمیم‌گیری منطقی انجام دهد. در واقع برای هر بازی، ۳ بخش زیر ضروری است:

- بازیکنان- استراتژی بازیکنان (در حیطه قواعد بازی). – دریافت‌ها و میزان بهینگی.

در هر بازی باید بازیکنان و عملکرد هر کدام از آن‌ها تعیین و مشخص شود. استراتژی بازیکنان برای سایر بازیکنان حریف نامشخص است و بر اساس احتمالات تعیین می‌شود. سپس بر اساس احتمالات، پیش‌بینی و درخت تصمیم ایجاد می‌شود. همچنین هر بازی تحت قواعدی مشخص می‌باشد و استراتژی بازیکنان و حریف، بر این اساس مشخص می‌شود. انتظار می‌رود بر اساس قواعد بازی، بازیکنان استراتژی‌های خود را چنان انتخاب کنند که بهترین نتیجه (بهینگی) حاصل شود.

در حالت کلی بازی‌ها را می‌توان به ۳ گروه زیر دسته‌بندی نمود که در شکل ۱. نشان داده شده است:

بازی‌های مهارتی – بازی‌های شانسی. – بازی‌های استراتژیک.



شکل ۱. دسته بندی انواع بازی‌ها

بازی‌های مهارتی، بازی‌های یک نفره‌ای هستند که در آن، بازیکن، کنترل تمام پیامدها را بر عهده دارد، مانند جلسه آزمون.

بازی‌های شانسی، بازی‌های یک نفره‌ای هستند که در آن بر خلاف بازی‌های مهارتی، بازیکن کنترل کاملی بر پیامدها ندارد و در واقع بخشی از رخدادهای بازی‌های شانسی، مبتنی بر انتخابِ

بازیکنان است و در واقع مبتنی بر عدم قطعیت می‌باشند. بازی‌های شانسی به دو گروه زیر تقسیم‌بندی می‌شوند:

-بازی شانسی شامل خطر: در این نوع بازی، بازیگر می‌تواند احتمالی را به هر یک از حرکت-های طبیعت و نتایج هر کدام از اعمال خود، اختصاص دهد. این گروه از بازی‌ها می‌توانند با استفاده از مفهوم امید ریاضی حل شوند. هنگامی که مقدار مطلوبیت مورد انتظار، در بازی‌های باخطر، به کمک بازیکن بیاید، نسبت به امید ریاضی، قاعده بهتری است. در واقع نظریه مطلوبیت فرض می‌کند که تصمیمات بر اساس این‌که چه نتایجی برای بازیکن و به جای تابع هدف خود، ارزشمند است، ساخته می‌شود. اگرچه بطور قطع در برخی مواقع، مقدار نتایج نهایی ساده، معادل با مقدار مطلوبیت است. علاوه بر این هر دو مفهوم دارای برخی زمینه‌های مشترک هستند. به عنوان نمونه، یک مقدار کوچک‌تر هرگز نمی‌تواند مطلوبیت یا نتیجه نهایی بزرگ‌تر از یک مقدار بزرگ‌تر از خود را داشته باشد.

- بازی‌های شانسی شامل عدم قطعیت: در این گروه از بازی‌ها، مانند بازی شامل خطر، یک بازیکن در مقابل طبیعت قرار می‌گیرد، اما بر خلاف بازی شامل خطر، بازیکن نمی‌تواند احتمالی را به حرکت‌های طبیعت اختصاص دهد. سه پیشنهاد برای تصمیم‌گیری در چنین حالتی وجود دارد: اصل بیشینه-بیشینه[1]. اصل بیشینه-کمینه[2]. اصل کمینه-بیشینه[3].

در اصل بیشینه-بیشینه، بازیکن استراتژی را اتخاذ می‌کند که دارای بیشترین نتیجه نهایی باشد. این رویکرد خیلی خوش‌بینانه بوده و قابل قیاس با استراتژی ریسک‌پذیری است.

اصل بیشینه-کمینه، در واقع اشار به این دارد که بازیکن از بدترین نتیجه نهایی ممکن اجتناب می‌کند. به عبارت دیگر بازیکن باید استراتژی را اتخاذ کند که بدترین سناریو را ارائه می‌کند و این یک رویکرد استراتژی خیلی بدبینانه ریسک گریز است و بدون توجه به بزرگی منفعت آن را نادیده می‌گیرد.

اصل کمینه-بیشینه در واقع اشاره به این دارد که بازیکن از استراتژی که بیشترین پشیمانی را به همراه دارد، اجتناب می‌کند و در واقع نقطه تعادل بین خیلی خوش‌بینانه و خیلی بدبینانه است.

بازی‌های استراتژیک، بازی‌های دو یا چند نفره‌ای هستند که در آن هر بازیکن، کنترل جزئی بر روی نتایج دارد. این نوع از بازی، به گروه‌های زیر دسته‌بندی می‌شوند:

---

[1] Max-Max
[2] Max-Min
[3] Min-Max

-بازی‌های ایستا و پویا[1]: در بازی‌های ایستا، حرکت بازیکنان بصورت ترتیبی می‌باشد، مانند شطرنج. اما در بازی‌های پویا، بازی با حرکت هم‌زمان چند بازیکنان انجام می‌شود و هیچ‌کدام از بازیکنان در مورد نحوه بازی حریف، اطلاعی ندارند.

-بازی‌های باهمکاری یا بدون همکاری[2]: در بازی‌های با همکاری، بازیکنان می‌توانند در هنگام بازی و یا قبل از آن، ارتباط آزاد داشته باشند، در صورتی که در بازی‌های بدون همکاری این امکان وجود ندارد. این نوع بازی‌ها دارای ساختار پیچیده‌ای هستند.

-بازی با اطلاع کامل و ناقص[3]: در بازی با اطلاعات کامل، تصمیم هر بازیکن مبتنی بر تمام حرکات قبلی سایر بازیکنان است، مانند شطرنج. اما اگر بازیکن به دلایلی، مجموعه‌ای از اطلاعات را در اختیار نداشته باشد، آن بازی با اطلاعات ناقص نامیده می‌شود.

-بازی با اطلاعات متقارن و نامتقارن[4]: در بازی با اطلاعات متقارن، هیچکدام از بازیکنان، اطلاعات بیشتری نسبت به سایرین ندارند و با جابه‌جایی استراتژی دو بازیکن، رخدادهای آن‌ها تغییر نمی‌کند، اما در بازی با اطلاعات نامتقارن، تعدادی از بازیکنان اطلاعات بیشتری نسبت به سایرین دارند.

-بازی‌های دونفره و چند نفره[5]: در بازی‌های دونفره تنها دو بازیکن مشغول رقابت هستند، اما در بازی‌های چند نفره، گروهی از افراد در حالِ رقابت بایکدیگر هستند.

-بازی‌های با مجموع صفر و مجموع غیر صفر[6]: در بازی‌های با مجموعِ صفر، ارزشِ بازی در فرآیندِ بازی، ثابت می‌ماند (کاهش یا افزایش نمی‌یابد) و سودِ یک بازیکن با زیانِ بازیکن دیگر همراه است. اما در بازی‌های مجموع غیر صفر، راهبردهایی موجود است که برای تمامی بازیکنان سودمند است.

-بازی‌های تصادفی و غیر تصادفی[7]: در بازی‌های غیر تصادفی، فرآیند بازی بصورتِ منطقی دنبال می‌شود اما در بازی‌های تصادفی، فرآیند بازی مبتنی بر تصادفی و اتفاقی می‌باشد.

-بازی‌های محدود و نامحدود[8]: در بازی‌های محدود، تعداد بازیکنان و استراتژی هر بازیکن محدود می‌باشد. اما در بازی‌های نامحدود بازیکنان استراتژی‌های نامحدودی دارند.

---

[1] Static and Dynamic Game Theory
[2] Cooperative and Non-Cooperative Game Theory
[3] Complete and Incomplete Information Game Theory
[4] Symmetric and Asymmetric Information Game Theory
[5] Two-person and N-person game
[6] Zero and Non-Zero-sum game
[7] Random and Non-Random game
[8] Finite and Non-Finite game

نمایش فضای بازی

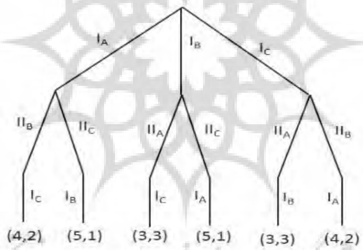
برای نمایش فضای بازی و حرکات بازیکنان، از سه روش زیر استفاده می‌شود[۱۲]:

شکل گسترده[1]. شکل متعارف[2]. شکل تابع مشخصه[3].

نمایش بازی در شکلِ گسترده، با استفاده از درخت، انجام می‌شود. در این روش، مجموعه گره‌های ممکن که بازیکن حرکتِ خود را از آن انتخاب می‌کند، مجموعه‌ی اطلاعاتی نامیده می‌شود.

به عنوان مثال، در یک بازی سه کارتِ (A,B,C) و دو بازیکن (I,II) وجود دارد. بازیکنِ Iیک کارت برمی‌دارد. سپس بازیکن II یکی از دو کارتِ باقیمانده را برمی‌دارد و کارت سوم به بازیکن اول داده می‌شود. بازیکنی که کارتِ A را داشته باشد سه واحد، دارنده‌ی کارتِ B دو واحد و دارنده‌ی کارتِ C یک واحد بدست می‌آورد. شکلِ (۱-۲) نمایشِ گسترده‌ی این بازی را نشان می‌دهد.

استراتژی هر بازیکن در این بازی به شکل گسترده، مشتمل بر مشخصاتِ کاملِ آن چه که وی در بازی انجام می‌دهد، می‌باشد. هر گره انشعابِ شاخه‌ها در درخت است. بنا بر شکلِ ۲ بازیکنِ I سه استراتژی دارد و بازیکنِ II دارای هشت استراتژی می‌باشد:



شکل ۲. نمایش گسترده‌ی بازی

S1: IA→IIB, IB→IIA, IC→IIA
S2: IA→IIB, IB→IIA, IC→IIB
S3: IA→IIB, IB→IIC, IC→IIA
S4: IA→IIB, IB→IIC, IC→IIB
S5: IA→IIC, IB→IIA, IC→IIA
S6: IA→IIC, IB→IIA, IC→IIB
S7: IA→IIC, IB→IIC, IC→IIA
S8: IA→IIC, IB→IIC, IC→IIB

[1] Extensive
[2] Normal
[3] Functional

نمایش بازی در شکل متعارف، شامل بازی‌های دو نفره بدون همکاری است که هر دو بازیکن بصورت همزمان، بازی می‌کنند. در این شکلِ نمایش، از ماتریس استفاده می‌شود. استراتژی بهینه‌ی یک بازیکن ممکن است خالص یا مرکب باشد. اگر استراتژی بهینه، تنها یکی از استراتژی‌های موجود باشد، به آن استراتژی بهینه خالص می‌گویند. اما اگر استراتژی بهینه مستلزم آن باشد که بازیکن تعدادی یا همه استراتژی‌های موجود را با احتمالاتی که به هر یک منسوب است استفاده کند، به آن استراتژی مرکب می‌گویند. اگر جمع دریافتی‌ها ثابت باشد، بازی را با جمعِ ثابت[1] می‌نامند. در حالت خاص اگر این مقدارِ ثابت مساوی صفر باشد، بازی را با جمعِ صفر[2] می‌نامند.

شکل تابع مشخصه، برای نمایش بازی‌های چند نفره مبتنی بر همکاری، مورد استفاده قرار می‌گیرد و حرکات بر اساس تابع مشخصه، نمایش داده می‌شوند.

انواع بازی‌ها

بازی‌های دونفره

این گروه از بازی‌ها به سه بخش زیر دسته بندی می‌شوند[۱۳]: -بازی‌های دو نفره جمع صفر. –بازی‌های دو نفره جمع غیر صفر. –بازی‌های اقتصادی.

بازی‌های دونفره جمع صفر

در این نوع از بازی‌ها، مجموع دریافتی بازیکنان (بهینگی) برابر صفر خواهد شد و هر آن چه که یک بازیکن بدست می‌آورد، توسط حریف از دست داده می‌شود[۱۴]. این گروه از بازی‌ها رقابتی هستند و مقدار بهینگی ساخته نشده و از بین نمی‌رود. در این گروه از بازی‌ها، علاق‌مندی دو بازیکن همیشه متضاد بوده و یک بازیکن با هزینه دیگری، برنده می‌شود (کارایی پارتو[3]). در واقع بازی‌های جمع صفر، به عنوان بازی‌های غیرمنصفانه شناخته می‌شوند. از این رو می‌توان این گروه از بازی‌ها را "بازی‌های جمع ثابت" نامید.

بازی‌های جمع صفر، به دو گروه محدود و نامحدود دسته‌بندی می‌شوند. بازی‌های جمع صفر نامحدود، بازی‌هایی هستند که حداقل یک بازیکن دارای شمار نامحدودی از استراتژی‌های خالص برای انتخاب را در اختیار دارد. تنها برخی از بازی‌های نامحدود قابل حل هستند، در حالی که بازی‌های محدود دارای حل می‌باشند. یک پاسخ برای بازی‌های جمع صفر، تعیین راهی است که

[1] Constant-sum-game
[2] Zero-sum
[3] Pareto-efficiency

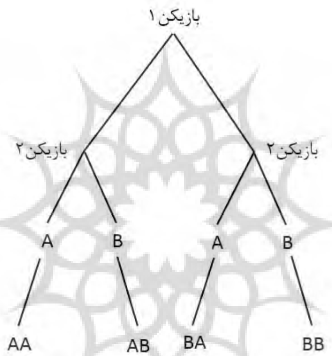
هر بازیکن بایستی حرکت کند. اگر هر دوبازیکن بر اساس این مشخصه حرکت کنند، در این صورت نتیجه نهایی حاصل شده، ارزش بازی[1] نامیده می‌شود.

نمایش بازی‌های دونفره جمع صفر محدود مبتنی بر ماتریس است. اما برای نمایش ارتباط میان درخت‌های بازی و تصمیم‌گیری، بازی‌ها را بصورت درختی نمایش می‌دهند.

ماتریس نمایش این بازی بصورت زیر است:

بازیکن ۲			
استراتژی **B**	استراتژی **A**	استراتژی	
AB	AA	استراتژی **A**	بازیکن۱
BB	BA	استراتژی **B**	

شکل ۳. بازی را بصورت درخت نمایش می‌دهد:



شکل ۳. نمودار درخت بازی برای بازی دونفره جمع صفر

جَیمز پی. بَیلی[2] و همکارانش یک مدل یادگیری چند عاملی مبتنی بر بازی‌های دونفره جمع صفر را ارائه نمودند[۱۵].

در اکثر بازی‌های جمع صفر بدون همکاری، استراتژی‌های خاصی وجود ندارد. بنابراین بازیکنان در صورت اطلاع از استراتژی انتخابی حریف سود می‌برند. در این نوع از بازی‌ها، به هر استراتژی، یک احتمال اختصاص داده می‌شود. همچنین تعیین استراتژی می‌تواند توسط برنامه‌ریزی خطی صورت پذیرد.

بازی‌های دونفره جمع غیرصفر

در بازی‌های دو نفره جمع صفر، سود یک بازیکن با زیان دیگری همراه است. این مساله برای بسیاری از بازی‌ها مخصوصا در حوزه اقتصاد و سیاست که در آن‌ها هر دو طرف می‌توانند تا

[1] Value of the game
[2] James P. Bailey

حدودی برنده یا بازنده باشند، محدود کننده است. بنابراین بازی‌های دونفره جمع غیر صفر[1] مطرح گردید[۱۶]. در این گروه از بازی‌ها، هر بازیکن، ماتریس نتایج خود را دارد. به عنوان مثال در مساله معمای زندانی، که در آن پلیس دو زندانی را بطور جداگانه مورد بازجویی قرار داده تا نتوانند در مورد جرم انجام شده همکاری کنند، طبق این گروه از بازی‌ها، می‌توان دو زندانی را به عنوان دو بازیکن در نظر گرفت که دو استراتژی اعتراف یا عدم اعتراف را دارند. ماتریس نتایج برای دو زندانی بصورت زیر می‌باشد:

جدول ۱- ماتریس نمایش برای دو زندانی

زندانی ۱یا۲	عدم‌اعتراف	اعتراف
اعتراف	(-۵و-۵)	(-۲۰و۰)
عدم‌اعتراف	(۰و-۲۰)	(-۱و-۱)

ماتریس‌های انفرادی برای دو بازیکن بصورت زیر می‌باشد:

$$A = \begin{bmatrix} -5 & 0 \\ -20 & -1 \end{bmatrix}, B = \begin{bmatrix} -5 & -20 \\ 0 & -1 \end{bmatrix}$$

بازی‌های چند نفره

در این گروه از بازی‌ها، تعداد بازیکنان بیشتر از ۳ می‌باشد و ساختار آن با بازی‌های دو نفره متفاوت می‌باشد و به بخش‌های زیر دسته بندی می‌شوند[۱۷]:

بازی‌های چند نفره بدون همکاری

در این گروه از بازی‌ها ائتلاف[2]هایی که توسط بعضی از بازیکنان برای کسب امتیاز نسبت به حریفان اختیار می‌شود صحبت می‌شود. راه حل عمومی برای بازی‌های چند نفره بدون همکاری، عموما بر پایه نقطه‌ی زینی[3] آن‌ها است. بدان معنا که هیچکدام از بازیکنان از نتیجه‌ی کسب شده تاسف نمی‌خورند، هنگامی که نتیجه‌ی سایر بازیکنان را مشاهده می‌کنند. تحلیل بازی‌های چند نفره عموما بسیار دشوار بوده و در نتیجه مفهومِ رفتار بهینه برای چنین موقعیت‌هایی قابل گسترش نیست. در این گروه از بازی‌ها، بازیکنان نمی‌توانند برای کسب نتایج بالاتر با یکدیگر همکاری کنند.

مفهوم تعادل استراتژیک[4] از این گروه می‌باشند. تعادل استراتژیک در واقع یک بازی چند نفره محدود می‌باشد که شامل یک مجموعه استراتژی متناهی غیرتهی $X_1, X_2, \dots, X_n$ و توابع مقدار

[1] Two-person-Non-Zero-Non-Cooperative
[2] Coalition
[3] Saddle points
[4] Strategic equilibrium

صحیحِ $u_1, u_2, \dots, u_n$ بر روی استراتژی‌ها می‌باشد. بردار استراتژی، تعادل استراتژی خالص[1] نامیده می‌شود، هر گاه داشته باشیم:

$$u_i(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) \geq u_i(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$$

مفهوم تعادل استراتژیک بیان می‌کند که یک گزینش خوب از استراتژی‌های انتخاب بازیکنان هنگامی که از بهترین پاسخ برای استراتژی انتخابی بازیکنان دیگر استفاده کند، یک تعادل استراتژی خالص را ایجاد خواهد نمود.

به عنوان مثال، دو بازیکن زیر:

$$(a)\ \begin{pmatrix} (3,3) & (0,0) \\ (0,0) & (5,5) \end{pmatrix}, \qquad (b)\ \begin{pmatrix} (3,3) & (4,3) \\ (3,4) & (5,5) \end{pmatrix}$$

در (a)، سطر اول و ستون اول که با $\langle 1,1 \rangle$ نمایش داده می‌شود، یک تعادل استراتژیک با نتیجه‌ی تعادلی $(3,3)$ است. اگر هر بازیکن معتقد باشد که دیگران قصد انتخاب استراتژی اول را دارند، در این صورت هیچ‌کدام از این دو بازیکن نمی‌خواهند به استراتژی دوم تغییر یابند. سطر و ستون دوم $\langle 2,2 \rangle$ نیز یک تعادل استراتژیک است. از آن‌جایی که نتیجه‌ی تعادلی آن $(5,5)$ است، در نتیجه هر دو بازیکن، این تعادل را ترجیح می‌دهند. در (b)، با توجه به تعریف، سطر اول و ستون اول $\langle 1,1 \rangle$ تعادل استراتژیک است. هیچکدام از بازیکنان نمی‌توانند با تغییر استراتژی سود کنند. به عبارت دیگر، هیچکدام از بازیکنان نمی‌توانند با تغییر صدمه یابند و اگر هر دو تغییر کنند، هر دو نتیجه‌ی بهتری را بدست خواهند آورد. بنابراین تعادل $\langle 1,1 \rangle$ به نسبت ناپایدار است.

مثال (a)نمونه‌ای از بازی است که در آن بازیکنان نتیجه‌ی مشابه‌ای را دریافت می‌کنند، اما اجازه‌ی برقراری ارتباط ندارند. اگر اجازه‌ی ارتباط داشته باشند اقدام مشترکی را که بیشترین نتیجه را به همراه دارد، انتخاب می‌کنند.

در یک بازی بدون همکاری، اگر بازیکنان اجازه‌ی ارتباط داشته باشند و به توافق غیر رسمی نیز برسند، می‌توان انتظار داشت که بازیکنان دارای یک تعادل استراتژیک باشند. از آن‌جایی که هیچ‌گونه توافق سازنده‌ای ایجاد نمی‌شود، توافقی انتظار می‌رود ایجاد شود که در واقع از تحمیل اراده‌ی خود بر دیگران[2] است. در این حالت هر بازیکن در برابر استراتژی که دیگران برای استفاده اعلان می‌کنند، مقدار درآمد خود را بیشینه می‌کند.

توسعه‌ی این روش برای آن‌که بازیکنان اجازه داشته باشند تا از استراتژی‌های ترکیبی استفاده کنند می‌تواند مفید باشد. یکی از مشکلات نظریه بازی بدون همکاری، این است که بطور معمول

[1] Pure strategic equilibrium
[2] Self-enforcing

برای بردارهای نتیجه‌ی متفاوت، تعادل بسیاری وجود دارد. یکی دیگر از مشکلات این است که اگر حتی یک تعادل استراتژیک یکتا هم وجود داشته باشد، ممکن است نتوان آن را به عنوان پاسخ منطقی و یا درآمد قابل پیش‌بینی در نظر گرفت.

به عنوان مثال بازی زیر را در نظر بگیرید:

$$\begin{pmatrix} (3,3) & (0,2) \\ (2,1) & (5,5) \end{pmatrix}$$

ماتریس نتایج متناظر بصورت زیر خواهد بود:

$$A = \begin{pmatrix} 3 & 0 \\ 2 & 5 \end{pmatrix}, B = \begin{pmatrix} 3 & 2 \\ 1 & 5 \end{pmatrix}$$

استراتژی ماکزیمم-مینیمم برای بازیکن ۱ برابر با $\left(1/_2 , 1/_2\right)$ و برای بازیکن ۲ برابر با $\left(3/_5 , 2/_5\right)$ است. سطح اطمینان برابر با $(V_I, V_{II}) = \left(51/_2 , 13/_5\right)$ خواهد بود.

در این مساله دو تعادل استراتژیک آشکار و متناظر با نتایج $(3,3)$ و $(5,5)$ وجود دارد. هر دو بازیکن دومین تعادل استراتژیک را ترجیح می‌دهند، زیرا به هر دو بازیکن به جای مقدار ۳ مقدار ۵ را می‌دهد. اگر آن دو با هم هماهنگ باشند این نتیجه قابل انتظار است. اما اگر نتوانند ارتباط برقرار کرده و اگر هر دو بازیکن فکر کنند که دیگری استراتژی اول را انتخاب می‌کند، آن‌گاه هر دو استراتژی اول را انتخاب کرده و هر دو نتیجه‌ی ۳ را دریافت می‌کنند.

بازی‌های چندنفره با انگیزه ترکیبی

یک بازی با n بازیکن را یک بازی چند نفره با انگیزه ترکیبی نامیده می‌شود، هرگاه:

هر بازیکن i یک مجموعه استراتژی محدود $S_i, i = 1, \dots, n$ و یک تابع بهینگی نتیجه $u_i \in S_1 \times s_2 \times \dots \times s_n \to \Re$ را داشته باشد و بصورت همزمان استراتژی $s_i \in S_i$ را انتخاب نموده و نتیجه u_i را دریافت کند.

بنابراین برای هر بازی بایستی استراتژی‌های مختلف و توابع بهینگی تعیین شوند و نتیجه‌ی هر بازیکن تابعی از تمام استراتژی‌ها بوده و فقط مربوط به یک بازیکن نیست. بنابراین تعادل نَش[1] برای یک بازی چند نفره با انگیزه ترکیبی به عنوان مجموعه‌ای از استراتژی‌ها بصورت زیر تعریف می‌شود:

$$u_i(S_{N1}, \dots, S_{Nn}) \geq u_i(s_1, \dots, s_n), \forall s_i \in S_i$$

مانند بازی‌های دو نفره، اگر تمامی بازیکنان قصد انجام عملی مشابه را داشته باشند، آن‌گاه هر بازیکن علاقه دارد تا یک استراتژی تعادل نَش را بازی کند. هر بازیکن باید درباره استراتژی

[1] Nash equilibrium

تعادل نَش آگاه باشد و همچنین بداند که بازیکنان دیگر درباره‌ی آن می‌دانند. همچنین این استراتژی باید توسط تمامی بازیکنان پذیرفته باشد و هیچ‌کدام از بازیکنان انگیزه‌ای برای انحراف از آن نداشته باشند. بنابراین پاسخ $S_{N1}, \dots, S_{Nn}$، یک تعادل نَش خواهد بود اگر:

$$\frac{\delta u_i}{\delta s_i}(S_{N1}, \dots, S_{Nn}) = 0$$

و اگر پاسخ یکتا داشته باشد، آن‌گاه این پاسخ تنها تعادل نَش برای بازی خواهد بود.

مثال حداکثر نمودن بودجه:

سازمان مدیریت و بودجه دارای سه منبع سرمایه‌گذاری است. درآمد سرانه مالیت‌های محلی و مطالبات کاربری که مرتبط با کمک‌های مردمی است (u_1)، سرانه‌ی یارانه‌های اضافی از استان‌ها به نواحی محروم (u_2) و بودجه‌ی اضافی از دولت برای خدمات خاص شهری (u_3) اختصاص می‌یابند.

سه فرمول بودجه برای شورای شهر عبارتند از:

$$u_1(a,b,c) = 36a + 5bc - a^2$$
$$u_2(a,b,c) = (a-3b)^{1/2} + c, a \geq 3b$$
$$u_3(a,b,c) = 4824c + 2ab - ac^2$$

بطوری که a جمعیت در هر شهر، b درصد جمعیت نواحی محروم و c تعداد ساعت‌های کاری است که برای فراهم نمودن تسیلات اضافی بطور متوسط در سه سال اخیر صرف شده است (بودجه هزینه).

می‌توان مشاهده کرد که:

$$1)\ \frac{\delta u_1}{\delta a} = -2a + 36$$
$$2)\ \frac{\delta u_2}{\delta b} = -3/2\,(a-3b)^{-1/2}$$
$$3)\ \frac{\delta u_3}{\delta c} = 4824 - 2ac$$

برابری هر کدام از این مشتقات جزئی با صفر، مقادیر زیر را خواهد داد. از معادله‌ی (1):

$$a = 18$$

با قرار دادن این مقدار در معادله‌ی (2):

$$0 = -3/2\,(18-3b)^{-1/2}$$
$$b = 6$$
$$c = 134$$

و با قرار دادن این مقدار در معادله‌ی (3):

مقادیر بودجه از این پاسخ‌های یکتا برابر با مقادیر زیر خواهند بود:

$u_1 = 4344$ **ریال**

$u_2 = 134$ **ریال**

$u_3 = 323424$ **ریال**

مشتقات دوم برابر هستند با:

$$\frac{\delta^2 u_1}{\delta a^2} = -2$$
$$\frac{\delta^2 u_2}{\delta b^2} = -{}^{9}\!/_{4}\,(a-3b)^{-3/2}$$
$$\frac{\delta^2 u_3}{\delta c^2} = -2a$$

بدیهی است از آن‌جایی که مقادیر a، b و c همیشه مثبت هستند، بنابراین هر سه مقدار بالا همیشه زیر صفر خواهند بود و بنابراین پاسخ موجود یک تعادل نش یکتا بوده و درآمد بودجه را بیشینه خواهد کرد.

بازی‌های چند نفره با همکاری جزئی

بازی‌های چند نفره با همکاری جزئی به بازی‌هایی گفته می‌شود که تشکیل ائتلاف در آن‌ها مجاز و حتی در بعضی موارد ضروری است. البته تا حدی که آن‌ها را به بازی‌های تمام همکاری کاهش ندهد. ویژگی این گروه از بازی‌ها، قدرت بوده و توسط بازیکنانی که آن را برای رسیدن به نتیجه‌ی مورد نظر خود به کار می‌برند، توزیع و مورد استفاده قرار می‌گیرد. این گروه از بازی‌ها با مفهوم یک تابع مشخصه همراه است که در واقع یک مقدار عددی را به هر ائتلاف امکان پذیر نسبت می‌دهد. مانوئل جیمز[1] و همکارانش یک مدل از این بازی، برای بازاریابی در شیلات ارائه دادند[۱۸].

مثال شهادت دادن:

یک دلال سهام شرکت و یک وکیل شرکت در بازرگانی داخلی شرکت مورد ظن بوده و در دو دفتر جداگانه توسط بازجویی از بخش کلاهبرداری پلیس پایتخت مورد سوال قرار می‌گیرند. مدارک و شواهدی که توسط همکاران ارائه شده، ضمنی بوده و برای محکوم کردن هر دو کافی

[1] Manuel Games

نیست، مگر آن‌که یکی علیه دیگری شهادت دهد. در نتیجه بازجویان در صورتی که هر کدام علیه دیگری شهادت دهند، به آن‌ها پیشنهاد رهایی و پاداش از طرف موسسات مالی را می‌دهند.

جدول ۱. ماتریس نتایج برای بازی شهادت دادن

وکیل	دلال سهام شرکت		
	استراتژی	خودداری از همکاری با بازجو	همکاری با بازجو
	خودداری از همکاری با بازجو	$(3,3)$	$(\mathbf{4},\mathbf{1})$
	همکاری با بازجو	$(4,1)$	$(\mathbf{2},\mathbf{2})$

اگر هر دو همکاری نکنند، در صورت عدم وجود هرگونه مدرکی برای اعمال مجازات‌های جدی، هر دو با سرزنش و توبیخ رسمی روبرو می‌شوند. اگر هر دو همکاری کنند، هر دو همیشه از تجارت معلق و از همکاری با شرکت محروم می‌شوند، اما زندانی نمی‌شوند. اگر یک از آن‌ها علیه دیگری شهادت دهد، دیگری زندانی شده و اولی به خاطر صداقت و همکاری بهترین نتیجه ممکن را بدست آورده و بر شهرت او برای این صداقت افزوده می‌شود.

این بازی دارای چهار ائتلاف ممکن است:

ائتلاف تهی که شامل هیچ بازیکنی نیست (C_0)، ائتلاف تکی که فقط شامل وکیل است (C_1)، ائتلاف تکی که شامل فقط یک دلال می‌باشد (C_2) و ائتلاف بزرگ که شامل هر دو بازیکن است (C_3). این موارد را می‌توان بصورت کلی زیر نمایش داد:

$$G = \sum_i C_i, i = 0,1,2,3$$

فرض کنید که ائتلاف C_i دارای تابع مشخصه‌ای (که در برخی موارد سطح امنیت نامیده می‌شود) است که بصورت $\omega\{C_i\}$ مشخص می‌شود، که بصورت حداقل نتیجه‌ای که ائتلاف C_i می‌تواند برای دسته‌های خود تضمین دهد، تعریف می‌شود. ارزش بازی برای ائتلاف تهی برابر با صفر است:

$$\omega\{C_i\} = 0$$

ائتلاف تکی که تنها شامل وکیل است، می‌تواند حداقل نتیجه ۲ را با انتخاب همکاری با بازجو تضمین کند. بنابراین سطح امنیت یا تابع مشخصه برای وکیل برابر است با:

$$\omega\{C_1\} = 2$$

بطور مشابه، ائتلاف تکی که تنها شامل دلال است، می‌تواند حداقل نتیجه ۲ را تضمین کند. بنابراین سطح امنیت یا تابع مشخصه برای دلال برابر است با:

$$\omega\{C_2\} = 2$$

در نهایت اگر هر دو ائتلاف بزرگ خود بصورت هماهنگ عمل کرده و با بازجو همکاری نکنند، می‌توانند نتیجه‌ای برابر با ۶ را تضمین نمایند:

$\omega\{C_3\} = 6$

از چهار تابع مشخصه داده شده، مشاهده می‌شود که اگر بازیکنان بخواهند حداکثر منفعت را داشته باشند، ائتلاف بزرگ را تشکیل خواهند داد.

بازی‌های چند نفره با همکاری کامل

یک بازی با همکاری کامل شامل دو عنصر است: یک مجموعه از بازیکنان و یک تابع مشخصه که ارزش ایجاد شده توسط زیر مجموعه‌های مختلف از بازیکنان در بازی را مشخص می‌کند. این مجموعه‌ی محدود از بازیکنان شامل n عنصر است و توابع مشخصه که با υ مشخص می‌شوند. همچنین تخصیص[۱]، مجموعه‌ای از اعداد است که بیانگر ارزش دریافتی توسط هر بازیکن است. ارزش شاپلی[۲] یکی از مهمترین راه حل‌های تخصیص ارزش می‌باشد. اِچ ژانگ[۳] و همکارانش بهبود سطح همکاری جمعیت از طریق مدل شناخت فردی را بر مبنای این گروه از بازی، پیشنهاد دادند[۱۹].

به عنوان مثال یک بازی با همکاری ساده به شرح زیر مطرح می‌شود:

سه بازیکن داریم، یعنی $N = \{1,2,3\}$. بازیکن ۱ فروشنده و بازیکن ۲ و۳ بصورت بالقوه خریدار می‌باشند. بازیکن ۱ یک واحد یکتا با قیمت ۴$ برای فروش دارد و حداکثر تمایل هر کدام از خریداران به خرید یک واحد است. بازیکن ۲ میل به پرداخت ۹$ برای محصول بازیکن ۱ دارد. در حالی که بازیکن ۳ میل به پرداخت ۱۱$ برای این محصول دارد. شکل ۴. این بازی را نشان می‌دهد.

شکل ۴. بازی مربوط به مثال

توابع مشخصه را برای این بازی بصورت زیر ایجاد می‌کنیم:

$V(\{1,2\}) = 9\$ - 4\$ = 5\$$
$V(\{1,3\}) = 11\$ - 4\$ = 7\$$
$V(\{2,3\}) = 0\$$
$V(\{1\}) = V(\{2\}) = V(\{3\}) = 0\$$
$V(\{1,2,3\}) = 7\$$

---

[1] Allocation
[2] Shapley
[3] H Zhang

این یک تعریف تا حدی شهودی برای تابع V است. اگر بازیکنان ۱ و ۲ باهم وارد معامله شوند، در این صورت مقدار منفعت کلی آن‌ها برابر با مقدار اختلاف بین میل به پرداخت خریدار و هزینه فروشنده یعنی $5\$$ است. همچنین اگر بازیگر ۱ و ۳ باهم وارد معامله شوند، در این صورت مقدار منفعت کلی آن‌ها برابر با مقدار اختلاف بین میل به پرداخت خریدار و هزینه فروشنده یعنی $7\$$ است. بازیکن ۲و ۳ نمی‌توانند با هم ارزشی را ایجاد کنند. زیرا هر دو به دنبال فروشنده هستند و نه خریدار. از آن‌جایی که بین هیچ بازیکن با خودش نمی‌تواند معامله‌ای رخ دهد، بنابراین هیچ بازیکنی نمی‌تواند برای خود ارزشی ایجاد کند. در نهایت توجه کنید که $V(\{1,2,3\})$ یک مجموعه معادل با $7\$$ بوده و نه $5\$ + 7\$ = 12\$$. دلیل این است که بازیکن ۱ تنها یک محصول برای فروش دارد و با این که دو خریدار در مجموعه $\{1,2,3\}$ وجود دارد، بازیکن ۱ تنها می‌تواند به تنهایی با یکی از آن‌ها معامله کند. در این حالت فرض می‌شود که بازیکن ۱ با بازیکنی که بیشترین میل به پرداخت را دارد، یعنی بازیکن ۳، معامله می‌کند و دلیل این امر آن است که مقدار $V(\{1,2,3\})$ به جای $5\$$ برابر با $7\$$ است.

بازی‌های تکراری

در شرایط تعامل‌های تکراری، بازیکنان یاد می‌گیرند که به گونه‌ای با استراتژی‌هل هماهنگ بشوند که نتایج ناکارآمد حاصل نشود. الگوی این گروه از بازی‌ها بر مبنای داد و ستد افراد در دنیا واقعی می‌باشد. بنابراین مفاهیمی مانند اعتبار، تهدید و کمال مطرح می‌شود.

این گروه از بازی‌ها به دو دسته‌ی بازی‌های با تکرار نامحدود و بازی‌های با تکرار متناهی دسته بندی می‌شوند[۲۰].

بازی‌های تکراری نامحدود

سرمایه‌گذاری تبلیغات نمونه‌ای از بازی‌های با تکرار نامحدود است. در این مثال اگر بخواهیم محرک‌هایی که دو شرکت بر صنعت بهداشت و درمان خصوصی انگلیس مسلط هستند، گروه عمومی بهداشت و درمان (GHG) و انجمن آینده نگر متحد بریتانیا $(BUPA)$. خدمات درمانی ملی (NHS) دولتی، از تسهیلات درمانی خصوصی برای جبران کسری در مقررات درمانی عمومی استفاده می‌کنند. به هرحال روابط دور از سادگی است. شرکت‌های درمانی خصوصی به هزینه درمانی ایالت به وسیله فروش با تخفیف کمک مالی می‌کنند، اما درگیری‌های آن‌ها و تبلیغات خوب اطراف آن، به وسیله افزایش مشتریان، منفعت را افزایش می‌دهد. بطور کلی منفعت هر شرکت بطور منفی مبتنی بر این است که شرکت‌های دیگر چه مقدار کمک مالی به تدارکات درمانی ملی (NHS) می‌کنند. با کمک مالی بیشتر (GHG)، جذب کمتری را برای $(BUPA)$

خواهد داشت و برعکس. اگر موضوعات با توجه به کمک به (NHS) بصورت زیاد یا کم تقسیم-بندی شوند، در این صورت نتایج بصورت جدولِ ۲. خواهد بود:

جدول ۲. ماتریس نتایج برای دو شرکت درمانی خصوصی

		BUPA	
GHG	استراتژی	کمک مالی زیاد برای **NHS**	کمک مالی کم برای **NHS**
	کمک مالی زیاد برای **NHS**	۲۰,۲۰	۴۰,۱۰
	کمک مالی کم برای **NHS**	۱۰,۴۰	۳۰,۳۰

اگر مانند مساله معمای زندانی این بازی فقط یک بار انجام شود، در این صورت با استفاده از استراتژی مینیمم-ماکزیمم یک تعادل نش در $(20,20)$ وجود خواهد داشت. با انتخاب یک استراتژی، استراتژی شرکت دیگر مشخص می‌شود و بنابراین هیچ شرکتی نمی‌تواند عملکرد خود را بهبود دهد. به هر حال این پاسخ غالب، بدتر از دیگر استراتژی که هر دو شرکت عمل مشابه را انجام دهند $(30,30)$، است و مساله برای رقابت شرکت‌ها این است که چگونه استراتژی‌های‌شان را روی نتیجه بهینه و بدون تثبیت قیمت هماهنگ کنند. در بازی‌های یک مرحله‌ای این امر امکان‌پذیر نیست، زیرا انگیزه‌ای آشکار برای کمک‌های مالی وجود دارد. با این حال اگر فعل و انفعال بین دو شرکت بصورت نامتناهی تکرار شود، این امکان برای هر دو شرکت وجود دارد که اعمال خود را روی نتیجه پارتو-موثر[1] هماهنگ کنند.

دو مفهوم کمک به توضیح این می‌کند که چگونه و چرا این امر اتفاق می‌افتد. یکی استراتژی تنبیه[2] است که در آن بازیکن یک استراتژی را صرفا بر اساس آن‌چه بازیکن دیگر انجام داده و به منظور تنبیه او به خاطر عدول از نتیجه کارای پارتو انتخاب می‌کند. استراتژی تنبیه، تنها اگر بخشی از تعادل نش کامل در زیربازی بی‌عیب باشد، می‌تواند موثر واقع شود. در این مثال وضعیت می‌تواند بدین صورت باشد که هر شرکت با استراتژی کمک مالی کم شروع کرده و تا زمانی که هیچ شخصی از وضعیت موجود منحرف نشود این آرایش می‌تواند تداوم داشته باشد. اگر هرکدام از شرکت‌ها استراتژی کمک مالی زیادی را اتخاذ کنند، در آن رویداد، شرکت رقیب تضمین می‌کند که پس از آن همیشه استراتژی کمک مالی زیاد را اتخاذ کند.

[1] Pareto-efficient
[2] Punishing strategy

این نوع خاص از استراتژی تنبیه به استراتژی محرک[1] معروف است که عمل یک بازیکن در بازی سبب می‌شود که بازیکن دیگر بطور مرتب به راه کار دیگری تغییر پیدا کند. در این مثال، اگر هر کدام از بازیکنان استراتژی کمک مالی زیاد را انتخاب کنند، یک دوره تنبیه نامحدود را در بر خواهد داشت. اگر یکی از شرکت‌ها سطح کمک‌هایش را افزایش دهد، شرکت دیگر تضمین می‌دهد که پس از آن بصورت مشابه عمل کند. شرکتی که اول استراتژی کمک مالی زیاد را اتخاذ می‌کند، منفعت را در اولین دوره از ۳۰ به ۴۰ افزایش خواهد داد، اما پس از آن سالیانه به ۲۰ نزول خواهد داشت و بازی در $(20,20)$ به تعادل خواهد رسید که یک نتیجه‌ی زیر بهینه برای هر دو طرف است.

برای یک استراتژی محرک به منظور استمرار نتیجه کارای-پرتو $(30,30)$ در این مثال، تنبیه و توافق برای استمرار نتیجه کارای-پارتو نباید مضحک باشد. در این مثال تهدید به تنبیه کاملا منطقی است، زیرا اگر یک شرکت استراتژی کمک مالی زیاد را انتخاب کند، از آن جایی که این حرکت افزایش منفعت بعدی شرکت را از ۱۰ به ۲۰ تضمین می‌کند، بنابراین برای شرکت دیگر نیز گزینش استراتژی کمک مالی بیشتر، منطقی خواهد بود. استراتژی تنبیه برابر با تعادل نش برای بازی‌های یک مرحله‌ای است.

تعهد برای استمرار موافقت ضمنی استراتژی کمک مالی کم، در این مثال نیز معتبر است. بطور کلی سازمان‌ها در بخش سود، به دنبال حداکثر کردن سود تنزیل یافته‌ی کلی هستند، بطوری که نتیجه همکاری در $(30,30)$، تا زمانی که ارزش فعلی همکاری بزرگ‌تر از ارزش فعلی منحرف شدن بوده و شرکت‌ها نتوانند در آینده بیشتر تخفیف دهند، بطور نامحدود برقرار خواهد بود. از آن جایی که یک بازی با تکرار نامحدود با زمان توسعه می‌یابد، بنابراین لازم است تا برخی نتایج آینده تا حدی تنزیل یابد. نتایج باز همان ارزش از دست می‌دهند، بنابراین مجموع پولی که در آینده دریافت می‌شود، باید امروز ارزش پایین‌تری داشته باشد و برعکس نتیجه‌ای که امروز دریافت می‌شود، بایستی ارزش بالاتری در آینده داشته باشد.

فرض کنید r نرخ بهره بالقوه باشد. بنابراین $d = 1/(1+r)$ نرخ تخفیف است. با این نرخ تخفیف، ارزش فعلی V_{now} برای ابقای استراتژی کمک مالی کم $V_{now}(small)$ بصورت زیر بیان می‌شود:

$$V_{now}(small) = 30 + 30d + 30d^2 + \cdots$$

بنابراین:

$$dV_{now}(small) = 30d + 30d^2 + 30d^3 + \cdots$$

[1] Trigger strategy

بنابراین:

$$(1-d)V_{now}(small) = 30$$

یا:

$$V_{now}(small) = {}^{30}\!/_{(1-d)}$$

ارزش فعلی انحراف از این نتیجه با همکاری و اتخاذ استراتژی کمک مالی زیاد، $V_{now}(large)$، بصورت زیر خواهد بود:

$$V_{now}(large) = 40 + 20d + 20d^2 + \cdots$$

بنابراین:

$$dV_{now}(large) = 40d + 20d^2 + 20d^3 + \cdots$$

بنابراین:

$$\begin{aligned}(1-d)V_{now}(large) &= 40 + 20d - 40d + 20d^2 - 20d^2 + \cdots \\ &= 40(1-d) + 20d\end{aligned}$$

بنابراین:

$$V_{now}(large) = 40 + {}^{20d}\!/_{(1-d)}$$

تا زمانی که $V_{now}(small) \geq V_{now}(large)$ باشد، این نتیجه با همکاری بصورت نامحدود باقی خواهد ماند. به عبارت دیگر اگر:

$$^{30}\!/_{(1-d)} \geq 40 + {}^{20d}\!/_{(1-d)}$$
$$d < 1, 30 \geq 40 - 20d$$
$$d \geq {}^{1}\!/_{2}$$

بنابراین در بازی تکراری نامحدود که در این مثال مطرح شد، هر دو شرکت خدمات درمانی تا زمانی که نرخ تخفیفی بزرگتر از ۰,۵ داشته باشند، تعادل کارای-پارتو در (30,30) باقی خواهد بود. اگر نرخ تخفیف کمتر از ۰,۵ باشد، هر شرکت از تعادل (30,30) خارج خواهد شد و بنابراین با محرک تنبیه داده شده، حفظ شود، زیرا تهدید آینده‌ی تنبیه به اندازه کافی بازدارنده نیست. بازیکنان اهمیت بسیار زیادی را به منفعت جاری در هزینه منفعت آینده می‌دهند و تعهد می‌دهند که حفظ ترتیب ضمنی بیشتر معتبر نیست.

با تعادل نش چندگانه، پیش‌گویی یکتایی برای آخرین تکرار وجود نخواهد داشت، بنابراین تهدیدی برای القای بازی با راه حلی با همکاری به بازیکنان دیگر ندارند.

بازی‌های تکاملی

الگوی این گروه از بازی‌ها مبتنی بر مباحث بیولوژیک است و برازندگی[1] را به عنوان یک فاکتور استراتژیک در پدیده‌های تکاملی معرفی می‌کند[۲۱]. در بازی‌های تکاملی دو رویکرد وجود دارد. در رویکرد نخست، مفهوم استراتژی پایدار تکاملی[2] به عنوان یک ابزار مهم در تحلیل استفاده می‌شود. در رویکرد دوم، در یک مدل ساده از فرآیند، تغییرات استراتژی بصورت پویا برای بررسی خواص تکاملی در جمعیت مورد مطالعه قرار می‌گیرد.

الگوی رویکرد نخست مبتنی بر مساله باز-کبوتر[3] می‌باشد که در آن دو موجود برای بدست آوردن منابع با ارزشِ V با یکدیگر رقابت می‌کنند. در این مساله، باز رفتار پرخاشگری را شروع می‌کند و تا زمانی که زخمی شود و یا یک از حریفان عقب نشینی نماید، متوقف نمی‌شود. اما کبوتر اگر یکی از حریفان رفتار پرخاشگری از خود نشان دهد، به سرعت عقب نشینی می‌کند. در این نوع از بازی، فرض‌های زیر در نظر گرفته می‌شود:

هنگامی که دو موجود رفتار پرخاشگری از خود نشان دهند، تعارض بین آن‌ها روی می‌دهد و هر یک با احتمال مساوی آسیب می‌بینند.

تعارض، ارزش مورد انتظار هر یک از بازیکنان را کاهش می‌دهد که با مقدار ثابت C نشان داده می‌شود.

هنگامی که یک باز با کبوتر برخورد می‌کند، کبوتر به سرعت عقب نشینی کرده و باز تمامی منابع را تصاحب می‌کند.

هنگامی که دو کبوتر با هم برخورد می‌کنند، منابع را بصورت مساوی بین خود تقسیم‌بندی می‌کنند. جدول ۴. بازده این بازی را نشان می‌دهد:

جدول ۴. ماتریس بازده بازی

	باز	کبوتر
باز	$\frac{1}{2}(V-C)$	$\boldsymbol{V}$
کبوتر	۰	$\boldsymbol{\frac{V}{2}}$

در استراتژی‌های پایدار تکاملی، باید خصوصیت‌هایی وجود داشته باشد که اگر هر بازیکن، بطور تقریبی آن را دنبال کند، تغییری برای ناحراف از استراتژی پایدار وجود نداشته باشد:

$$\boldsymbol{\Delta}F(S_1, S_2)$$

---

1 Fitness
2 Evolutionarily Stable Strategy
3 Hawk-Dove

و به معنای تغییر در ارزش بازیکنی[۱] که استراتژی S_1 را برگزیده است، زمانی که بازیکن دیگر، استراتژی S_2 را دنبال می‌کند. علاوه بر این هر بازیکن در جمعیت، یک ارزش اولیه[۲] دارد که با F_0 نشان داده می‌شود. در این صورت اگر $\boldsymbol{\delta}$ استراتژی تکاملی پایدار باشد و μ استراتژی تهاجم و تغییر باشد، آن‌گاه:

$$F(\delta) = F_0 + (1-p)\Delta F(\delta,\delta) + p\Delta F(\delta, C)$$
$$F(\mu) = F_0 + (1-p)\Delta F(\mu,\delta) + p\Delta F(\mu,\mu)$$

که در آن p نسبتی از جمعیت است که استراتژی تهاجم μ را دنبال می‌کند. از آن‌جایی که δ استراتژی پایدار تکاملی است، بنابراین ارزش برای هر بازیکنی که این استراتژی را دنبال می‌کند بایستی از ارزش بازیکن دیگر که استراتژی μ را دنبال می‌کند، بزرگ‌تر باشد $\left(\mathrm{F}(\delta) > \mathrm{F}(\mu)\right)$. زیرا در غیر اینصورت، بازیکنی که استراتژی μ را دنبال می‌کند از استراتژی پایدار تکاملی فاصله می‌گیرد.

در رویکرد دوم، تغییرات استراتژی بصورت پویا برای بررسی خواص تکاملی در جمعیت مورد مطالعه قرار می‌گیرد.

در سال ۲۰۱۸، جانتان نیوتن[۳] به بررسی بازی‌های تکاملی پرداخت[۲۲].

برای توضیح این رویکرد، مساله معمای زندانی را در نظر می‌گیریم. در این بازی بازیکنان یکی از استراتژی‌های همکاری[۴] یا عدم همکاری[۵] را دنبال می‌کنند. ماتریس بازده برای این مساله بصورت جدول ۵. می‌باشد:

جدول ۵.- ماتریس بازده بازی معمای زندانی

	همکاری	عدم همکاری
همکاری	(R, R')	$(\boldsymbol{S}, \boldsymbol{T'})$
عدم همکاری	(T, S')	$(\boldsymbol{P}, \boldsymbol{P'})$

در این جدول $T > R > P > S$ و $T' > R' > P' > S'$ است. همچنین فرض می‌شود که در این مساله، منفعت برای تمام افراد در جمعیت یکسان می‌باشد.

برای انجام این بازی بصورت تکراری، فرض‌های زیر را در نظر می‌گیریم:

[1] Individual Fitness
[2] Initial Fitness
[3] Jonathan Newton
[4] Cooperate
[5] Defect

در ابتدا فرض می‌کنیم که اندازه‌ی جمعیت به قدر کافی بزرگ است. در این صورت وضعیت جمعیت به وسیله نسبتی از جمعیت که استراتژی همکاری یا عدم همکاری را دنبال می‌کنند، قابل ارائه می‌باشد.

P_c نسبتی از جمعیت است که استراتژی همکاری را دنبال می‌کند.

P_d نسبتی از جمعیت است که استراتژی عدم همکاری را دنبال می‌کند.

W_c متوسط ارزش برای کسانی که استراتژی‌های همکاری را دنبال نمایند.

W_D متوسط ارزش برای کسانی است که استراتژی عدم همکاری را دنبال می‌کنند.

$\bar{W}$ متوسط ارزش برای کل جمعیت.

در این صورت مقادیرِ W_c، W_D و $\bar{W}$ می‌توانند بصورت زیر محاسبه شوند:

$$W_c = F_0 + P_c \Delta F(C,C) + P_d \Delta F(C,D)$$
$$W_D = F_0 + P_c \Delta F(D,C) + P_d \Delta F(D,D)$$
$$\bar{W} = P_c + P_d W_D$$

دوم، فرض می‌کنیم نسبتی از افراد که در جمعیت بعدی استراتژی‌های همکاری یا عدم همکاری را دنبال می‌کنند با نسبتی از افراد که در جمعیت فعلی استراتژی‌های همکاری یا عدم همکاری را دنبال می‌کنند، بصورت زیر مرتبط باشند:

$$P_c' = P_c \frac{W_c}{\bar{W}} , P_d' = P_d \frac{W_D}{\bar{W}}$$

این عبارت‌ها را می‌توانیم بصورت زیر بنویسیم:

$$P_c' - P_c = \frac{P_c(W_c - \bar{W})}{\bar{W}}$$
$$P_d' - P_d = \frac{P_d(W_D - \bar{W})}{\bar{W}}$$

در صورتی که فرض شود تغییر در استراتژی از یک نسل به نسل بعدی (در تکرار) کوچک باشد، آن‌گاه معادله‌های فوق می‌توانند بصورت معادلات دیفرانسیل زیر بیان شوند:

$$\frac{dp_c}{dt} = \frac{P_c(W_c - \bar{W})}{\bar{W}}$$
$$\frac{dp_d}{dt} = \frac{P_d(W_D - \bar{W})}{\bar{W}}$$

این معادله‌ها شرایط دینامیک را در نظریه‌ی بازی‌ها در حالت تکاملی تامین می‌کند و به عنوان تکرار دینامیکی[1] نامیده شود. در صورتی که شرایط تکرار دینامیکی برای مساله معمای زندانی به کار گرفته شود مقدار ارزش برای هر یک از استراتژی‌های بازی بصورت زیر خواهد بود:

$$W_c = F_0 + P_c \Delta F(C,C) + P_d \Delta F(C,D) = F_0 + P_c R + P_d$$

و همچنین:

$$W_D = F_0 + P_c \Delta F(D,C) + P_d \Delta F(D,D) = F_0 + P_c T + P_d P$$

و از آن جایی که $T > R$ و $P > S$ می‌باشد، بنابراین $W_D > W_c$ است و همچنین $W_D > \bar{W} > W_c$ خواهد بود، که در این صورت خواهیم داشت:

$$\frac{(W_c - \bar{W})}{\bar{W}} < 0$$

$$\frac{P_d(W_D - \bar{W})}{\bar{W}} > 0$$

در این صورت استراتژی‌های همکاری و عدم همکاری در جمعیت بعدی بصورت زیر خواهد بود:

$$P'_d = P_d . \frac{(W_D - \bar{W})}{\bar{W}}$$

$$P'_c = P_c . \frac{(W_c - \bar{W})}{\bar{W}}$$

ملاحظه می‌شود که نسبت جمعیتی که استراتژی همکاری را انتخاب می‌کنند در طول زمان به تدریج رو به کاهش می‌گذارد.

شکل ۵. حالت تکرار دینامیکی را برای مساله معمای زندانی بصورت یک دیاگرام فضای حالت[2] نشان می‌دهد:

شکل ۵. بازی مربوط به مثال

در این دیاگرام، نقطه‌ی سمت چپ موقعیتی از جمعیت را نشان می‌دهد که هر عضو از جمعیت از استراتژی عدم همکاری استفاده می‌کند و نقطه‌ی سمت راست در دیاگرام، موقعیتی را نشان می‌دهد که هر عضو از جمعیت از استراتژی همکاری استفاده می‌کند. نقاط میانی دیاگرام، وضعیت‌هایی را نشان می‌دهند که بخشی از جمعیت استراتژی عدم همکاری و بقیه استراتژی همکاری دارند. (هر حالتی از جمعیت می‌تواند با نقطه‌ی روی دیاگرام مرتبط باشد. به گونه‌ای که

[1] Replicator dynamics
[2] State-space-diagram

هرگاه $N\%$ از جمعیت، استراتژی عدم همکاری دارند، معادل نقطه‌ای از دیاگرام است که $N\%$ از خط بین دو استراتژی تا نقطه‌ی سمت چپ دارد)

پیکان‌های روی دیاگرام، مسیر تکاملی که جمعیت در طول زمان دنبال می‌کند را نشان می‌دهد. همچنین دایره توخالی سمت راست، مشخص کننده‌ی حالتی است که در آن هر فردی که استراتژی همکاری دارد، در واقع یک تعادل ناپایدار را ارائه می‌کند، زیرا در این حالت اگر جزء کوچکی از جمعیت از استراتژی همکاری منحرف شود، آن‌گاه شرایط تکامل دینامیک[1] جمعیت را از حالت تعادل منحرف می‌کند.

دایره‌ی توپُر سمت راست نیز مشخص کننده‌ی حالتی است که در آن هر فردی که استراتژی عدم همکاری دارد، یک تعادل پایدار را مشخص می‌کند زیرا در این حالت اگر جزء کوچکی از جمعیت از استراتژی عدم همکاری منحرف شود، آن‌گاه شرایط تکامل دینامیک جمعیت را به حالت تعادل اولیه برمی‌گرداند.

به راحتی ملاحظه می‌شود که حالت استراتژی پایدار تکاملی در مساله معمای زندانی در وضعیتی رخ دهد که هر فرد استراتژی عدم همکاری دارد و از آن جایی که این وضعیت تنها در حالت تعادل پایدار است که تحت شرایط تکرار دینامیکی برقرار است، بنابراین می‌توان بیان نمود که دو رویکرد با هم تطبیق دارند. باید توجه داشت که اگر مدل دینامیکی تکاملی به وسیله‌ی تکرار دینامیکی برقرار نشود، در آن صورت ارتباط بین حالت استراتژی تکاملی و حالت پایدار ضعیف خواهد بود.

هنگامی که ماتریس بازده مقادیرِ $(T = 2.8, R = 1.1, P = 0.1, S = 0)$ را برای جمعیت دارند، این مدل در دینامیک تکاملی با تکرار دینامیکی به وضعیتی که در آن هر فرد استراتژی عدم همکاری را دنبال می‌کند، همگرا می‌شود.

هنگامی که ماتریس بازده مقادیرِ $(T = 1.2, R = 1.1, P = 0.1, S = 0)$ را برای جمعیت دارند، این مدل در دینامیک تکاملی جمعیت را به یک چرخه‌ی نوسانی پایدار بین دو حالت می‌برد. در این چرخه، هر دو استراتژی همکاری و عدم همکاری در نواحی که بطور نوسانی بین دو استراتژی جا به جا می‌شود، وجود دارد.

دقت کنید که با این مقادیر خاص برای ماتریس بازده، حالت تکاملِ دینامیک با حالت تکرار دینامیکی تفاوت دارد. بنابراین با این مقدار بازده بین تکرار دینامیکی و استراتژی پایدار تکاملی شباهتی وجود ندارد.

یک پدیده‌ی مورد توجه زمانی رخ می‌دهد که ماتریس بازده مقادیرِ

---

[1] Evolutionary dynamic

$(T = 1.61, R = 1.01, P = 0.01, S = 0)$

را داشته باشد. با این مقادیر نواحی که بطور غالب به وسیله استراتژی همکاری اشغال شده ممکن است به وسیله‌ی استراتژی عدم همکاری مورد تهاجم قرار گیرد و نواحی که بطور غالب به وسیله‌ی استراتژی عدم همکاری اشغال شده است ممکن است به وسیله استراتژی همکاری مورد تهاجم قرار گیرد. بنابراین در این مدل استراتژی پایدار، در حالت دینامیک، نسبتی وجود ندارد.

بازی‌های جمعیتی

یک ایده بسیار مهم که با توجه به استراتژی تکاملی پایدار معرفی شد، این است که بازیکنان در نهایت استراتژی‌هایی را انتخاب می‌کنند که نتیجه‌ای بهتر از حد متوسط را ایجاد کند[۲۳]. مفاهیم کلیدی در این بخش عبارت‌هایی مانند "درنهایت[1]" و "بهتر از حد متوسط[2]" هستند. عبارت "درنهایت" به معنای وابستگی به زمان و یک محدودیت به عنوان گذر زمان است. بنابراین تلاش برای مدل کردن این که با معرفی سوال وابسته به زمان و دادن مجوز برای $\boldsymbol{t} \to \infty$، چه اتفاقی خواهد، طبیعی است.

در این جا یک مقدمه وجود دارد، $\boldsymbol{N}$ عضو از جمعیت وجود دارد و به طور تصادفی، دو عضو از این جمعیت برای بازی در مقابل هم انتخاب می‌شوند. فرض می‌کنیم که $\boldsymbol{N}$ یک عدد بسیار بزرگ باشد. بنابراین احتمال رویارویی دو عضو یکسان در عمل صفر است.

فرض می‌کنیم که بازی آن‌ها یک بازی دو نفره متقارن باشد. تقارن[3] در این جا به معنای $\boldsymbol{B = A^T}$ است، بدان معنا که بازیکنان می‌توانند بدون تغییر، نقش خود را تعویض کنند. در یک بازی متقارن، مهم نیست که کدام بازیکن شماره ۱ و کدام بازیکن شماره ۲ است.

فرض کنید که بازیکنان بسیاری در جمعیت وجود دارد. هر بازیکن که از جمعیت انتخاب می‌شود، یک استراتژی ترکیبی را انتخاب کرده و با ماتریس $\boldsymbol{A}$ در مقابل هر بازیکن دیگری که انتخاب می‌شود بازی می‌کند. این بازی یک رقابت تصادفی[4] نامیده می‌شود. توجه داشته باشید که همه بازیکنان از ماتریس نتایج یکسان استفاده خواهند کرد.

ممکن است بازیکنان در جمعیت از استراتژی‌های خالص $\boldsymbol{1, 2, \ldots, n}$ استفاده کنند. درصد بازیکنان جمعیت که از استراتژی j-ام استفاده می‌کنند برابر است با:

[1] Eventually
[2] Better-than-average
[3] Symmetry
[4] Random contest

$$p(player\ uses\ j) = p_j, p_j \geq 0, \sum_{j=1}^{n} p_j = 1$$

مجموعه‌ی $\boldsymbol{\pi}$ بصورت $\boldsymbol{\pi} = (\boldsymbol{p_1}, \dots, \boldsymbol{p_n})$ است. این عناصر $\boldsymbol{p_i}$ از مجموعه‌ی $\boldsymbol{\pi}$ فرکانس نامیده شده و بیانگر احتمال یک انتخاب ناگهانی منحصر به فرد در جمعیت است، که از استراتژی i-ام استفاده خواهد کرد. با رابطه زیر بیان می‌شود:

$$\prod = \left\{ \pi = (p_1, \dots, p_n) | p_j \geq 0, j = 1,2, \dots, n, \sum_{j=1}^{n} p_j = 1 \right\}$$

که مجموعه همه فرکانس‌های ممکن است. اگر دو بازیکن از جمعیت انتخاب شوند و بازیکن ۱، استراتژی i، و بازیکن ۲، استراتژی j را انتخاب کند. ما نتایج را از ماتریس A محاسبه می‌کنیم. برازندگی بازیکنی که استراتژی $\boldsymbol{1, 2, \dots, n}$ را بازی می‌کند، بصورت زیر تعریف می‌کنیم:

$$E(i, \pi) \sum_{k=1}^{n} a_{i,k} p_k = A\pi$$

این نتیجه مورد انتظار رقابت تصادفی است برای بازیکن ۱ که استراتژی i-ام را در مقایل استراتژی‌های ممکن $\boldsymbol{1, 2, \dots, n}$ استفاده کرده و با احتمال‌های $\boldsymbol{p_1, \dots, p_n}$ بازی می‌کند و ارزش استراتژی i-ام را در جمعیت اندازه‌گیری می‌کند. مشاهده می‌کنید که $\boldsymbol{\pi}$ شبیه یک استراتژی ترکیبی بوده و ما $\boldsymbol{E(i, \pi)}$ را به عنوان نتیجه مورد انتظار بازیکن ۱ بدین صورت که اگر بازیکن ۱ استراتژی خالص i و طرف مقابل (بازیکن ۲) استراتژی ترکیبی $\boldsymbol{\pi}$ را مورد استفاده قرار دهد، می‌شناسیم.

اکنون برازندگی مورد انتظار کل جمعیت بصورت زیر محاسبه می‌شود:

$$E(\pi, \pi) \sum_{i=1}^{n} p_i \left[\sum_{k=1}^{n} a_{i,k} p_k \right] = \pi A \pi^T$$

اکنون فرض کنید که فرکان$\boldsymbol{\pi} = (\boldsymbol{p_1}, \dots, \boldsymbol{p_n}) = \boldsymbol{\pi(t)} \in \boldsymbol{\Pi}$ س بتواند با زمان تغییر کند. بنابراین نیاز به مدلی داریم که تا شرح دهد که چطور فرکانس‌ها می‌توانند با زمان تغییر کنند و ما چگونه می‌توانیم از پویایی فرکانس به عنوان سیستم معادله دیفرانسیل زیر استفاده کنیم:

$$\frac{dp_i(t)}{d(t)} = p_i(t)\left[E(i,\pi) - E\big(\pi(t),\pi(t)\big)\right]$$
$$= p_i(t)\left[\sum_{k=1}^{n} a_{i,k}p_k - \pi(t)A\pi(t)^T\right], i = 1,2,\dots,n$$

یا بطور معادل:

$$\frac{dp_i(t)}{d(t)} = \left[E(i,\pi) - E\big(\pi(t),\pi(t)\big)\right]dt$$

که پویایی رونوشت ساز[1] نیز نامیده می‌شود.

ایده این است که نرخ رشد، در کدام درصد جمعیت استفاده کننده از استراتژی i به وسیله چه مقدار بزرگتر یا کمتر نتیجه مورد انتظار استفاده کننده استراتژی i ارزیابی شده و با برازندگی مورد انتظار استفاده کننده کل استراتژی‌ها، در جمعیت مقایسه می‌شود. استراتژی‌های بهتر باید با فرکانس افزایشی و بدترین استراتژی‌ها با فرکانس کاهشی استفاده شوند. در مورد یک بعدی، اگر برازندگی مورد استفادع در استراتژی خالص i بهتر از مقدار متوسط باشد، سمت راست مثبت خواهد بود و در صورت بدتر بودن از مقدار متوسط، مشتقِ $\boldsymbol{\frac{dp_i(t)}{d(t)}}$ مثبت بوده و سبب می‌شود که $\boldsymbol{p_i(t)}$ با پیشرفت زمان، افزایش یابد و استراتژی i با فرکانس افزایشی مورد استفاده واقع خواهد شد.

در این نقطه، شرایط اولیه $\boldsymbol{\pi(0) = \big(p_1(0),\dots,p_n(0)\big)}$ را مشخص نمی‌کنیم. اما می‌دانیم که $\boldsymbol{\sum_i p_i(0) = 1}$ است. همچنین باید پاسخ حقیقی معادله برای همه مقادیر $\boldsymbol{t > 0}$، بصورت $\boldsymbol{0 \le p_i(t) \le 1}$ و $\boldsymbol{\sum_i p_i(t) = 1}$ باشد.

بازی‌های بر ضد طبیعت

نظریه بازی‌ها حالت خاصی از نظریه تصمیم‌گیری است، یعنی نوعی تصمیم‌گیری در حالت عدم قطعیت که در آن یک یا چند نفر انسان صاحب عقل در مقابل تصمیم‌هایی که گرفته می‌شود، پاسخ نشان می‌دهند. در واقع این همان تصمیم‌گیری است با این تفاوت که حالت‌های طبیعی دیگر صرفا همراه با عدم قطعیت نبوده و در آن اختیار فرد یا افراد دیگری وجود دارد. بنابراین نظریه بازی‌ها خود حالتی از تصمیم‌گیری می‌باشد، از طرف دیگر تصمیم‌گیری حالت خاصی از نظریه بازی به شمار می‌آید که پاسخ بازیکنان بر ضد طبیعت بدون عقل می‌باشد. بازی‌های بر ضد طبیعت بازی‌هایی بر علیه یک بازیکن است که منفعت آن در بازی معلوم نیست و در خصوص احتمالاتی که استراتژی‌های خود را انتخاب می‌کنند، جهل کامل وجود دارد. به

[1] Replicator dynamics

عنوان مثال فردی که با یک ماشین بازیگر بازی می‌کند، دارای دو استراتژی است، استراتژی اول بازی کند و استراتژی دوم بازی نکند و منفعتی که ممکن است از ترکیب حالت‌های ممکن ماشین بدست آید، معلوم است. برای بعضی از این ترکیب‌ها، بازیگر منفعت مثبت دارد و برای برخی دیگر منفعت منفی دارد و اگر بازی نکند منفعت آن صفر است و احتمال آن که یک ترکیب ظاهر شود برای یک فرد معمولی که کارشناس احتمالات نباشد، نامعلوم است. بنابراین چنین بازی را می‌توان بر ضد طبیعت نامید.

کاربردهای موردی: تئوری بازی در ارتباط پهپادها مبتنی بر شبکه‌های بی‌سیم

تئوری بازی یک عنصر پایه در ارتباط پهپادها مبتنی بر شبکه‌های بی‌سیم می‌باشد که در مدیریت منابع[۲۴-۲۷]، مسیریابی پهپادها[۲۸-۳۰]، موقعیت‌یابی[۳۱]، مدیریت تداخل[۳۲] و امنیت[۳۳, ۳۴] مورد استفاده قرار گرفته‌اند. در موقعیت یابی، ارتفاع و زاویه ارتفاع یک پهپاد، بر عملکرد پوشش آن و قابلیت اطمینان پیوند در یک منطقه خدماتی تأثیر می‌گذارد. علاوه بر این، چگالی بهینه پهپادها در یک منطقه تابع محدودیت‌های ایمنی و تداخل است. تحقیقات مربوط به موقعیت‌یابی پهپاد بر به حداکثر رساندن پوشش سیستم و در عین حال به حداقل رساندن تداخل تمرکز دارد. در برنامه‌ریزی مسیر، با توجه به محدودیت انرژی، مسیر پهپادها در یک شبکه باید بهینه شود و کیفیت پیوند، تداخل و اجتناب از برخورد، بایستی در نظر گرفته شود. در امنیت، پارازیت و استراق سمع بین پهپادها و دستگاه‌ها دو مشکل اصلی امنیتی می‌باشند. هر دو ضررهای اقتصادی و سیاسی زیادی را به شرکت‌ها و کاربران وارد می‌نمایند. در مدیریت منابع، دستگاه‌های تلفن همراه و دستگاه‌های اینترنت اشیاء عمر باتری محدود و قابلیت ذخیره‌سازی محدود دارند. در نتیجه، در یک شبکه سلولی پهپاد، نیاز به پشتیبانی از حافظه نهانِ داده و انتقال محتوا می‌باشد. به هر پهپاد ممکن است وظایف متفاوتی (در حافظه پنهان یا رله) اختصاص داده شود و همچنین ممکن است کاربران مختلفی را برای خدمت انتخاب کند. هدف مدیریت منابع به حداکثر رساندن درآمد اپراتور(ها) از طریق بهینه‌سازی تکالیف وظایف و انتخاب کاربر است. علاوه بر این، اگر پهپادها متعلق به اپراتورهای مختلف هستند، رقابت بین اپراتورها نیز باید در نظر گرفته شود. تداخل هم در شبکه‌های سنتی زمینی و هم در شبکه‌های به کمک پهپاد وجود دارد. در مورد دوم، تداخل از سه منبع ناشی می شود: سایر پهپادهای ارتباطی، کاربران تلفن همراه و ایستگاه های کنترل زمینی.

استفاده از بازی‌های مبتنی بر همکاری در ارتباط پهپادها مبتنی بر شبکه‌های بی‌سیم

پهپادها می‌توانند دسترسی به طور مشترک به شبکه را برای کاربران به منظور کاهش تلفات آن‌ها فراهم کنند. در واقع پهپادهای یک اپراتور می‌توانند به طور مشترک وظایفی (مانند

مسیریابی، جمع‌آوری داده‌ها و غیره) را انجام دهند تا درآمد اپراتور را به حداکثر برسانند. بنابراین، نظریه بازی مبتنی بر همکاری ممکن است برای مدل‌سازی مسائل تخصیص نرخ، انتقال مشارکتی، ارسال بسته و غیره استفاده شود. اما اشکال بازی مبتنی بر همکاری این است که برای آن‌ها هیچ الگوریتم شناخته شده قابل اجرا در زمان چند جمله‌ای وجود ندارد، یعنی با افزایش اندازه شبکه ارتباطی، پیچیدگی زمانی افزایش می‌یابد. بنابراین، الگوریتم‌های اکتشافی متعددی معمولاً برای یافتن یک راه‌حل تقریباً بهینه در شبکه‌های ارتباطی بزرگ استفاده می‌شوند.

بازی تشکیل ائتلاف لذت بخش[1] یک کلاس ویژه از بازی‌های تشکیل ائتلاف است که در آن بازیکنان به نفع خود هستند و فقط به هویت بازیکنان در ائتلاف خود اهمیت می‌دهند و هر بازیکن نسبت به ائتلاف‌های مختلف رتبه ترجیحی دارد. در مرجع[۳۵]، تعدادی از پهپادها برای جمع‌آوری داده‌های مبتنی بر چندین وظیفه، با مکان‌یابی دلخواه در یک شبکه بی‌سیم موردی[2] پرواز مبتنی بر پهپاد مورد نیاز هستند. یک بازی تشکیل ائتلاف لذت‌گرا برای مدل‌سازی تعاملات بین پهپادها و وظایف به منظور تشکیل ائتلاف‌های غیرمرتبط استفاده می‌شود. هم وظایف و هم پهپادها، بازیکنانی هستند که بر اساس بازدهی خود تصمیم به پیوستن یا ترک یک ائتلاف می کنند. مجموع مطلوبیت هر ائتلاف با استفاده از یک تابع ارزش ائتلافی که به عنوان نسبت مقداری توان خروجی و تاخیر تعیین شده است، ارزیابی می شود. هر ائتلاف تشکیل شده به عنوان یک سیستم نظرسنجی متشکل از تعدادی پهپاد که بین وظایف مختلف برای جمع آوری و ارسال بسته ها به یک گیرنده مشترک حرکت می کنند، مدل سازی‌می شود.

با توجه به پیچیدگی محاسباتی، پهپادها بر اساس نزدیک‌ترین مسیر همسایه عمل می‌کنند. ائتلاف تا زمانی که به یک پارتیشن شبکه پایدار نَش برسد به روز رسانی را ادامه می‌دهد. نویسندگان عملکرد این الگوریتم را با الگوریتمی مقایسه کردند که وظایف را به طور مساوی در بین پهپادها تعیین می‌کند. نتایج شبیه‌سازی نشان می‌دهد که الگوریتم پیشنهادی از تخصیص مساوی از نظر میانگین بازده حداقل ۳۰ درصد بدون توجه به تعداد کارها بهتر عمل می‌کند.

استفاده از بازی‌های مبتنی بر عدم همکاری در ارتباط پهپادها مبتنی بر شبکه‌های بی‌سیم

برخلاف بازی‌های مشارکتی، تئوری بازی‌های غیرهمکاری به سناریویی می‌پردازد که بازیکنان برای به حداکثر رساندن سود خود با یکدیگر رقابت می‌کنند. بنابراین، این نوع بازی فرض می‌کند که همه بازیکنان به نفع خود عمل می‌نمایند.

---

1. Hedonic coalitional formation game
2. Ad-hoc

انواع مختلفی از بازی‌های غیرهمکاری مانند بازی‌های دیفرانسیل، بازی‌های بیزی و غیره وجود دارد. تئوری بازی‌های مبتنی بر عدم همکاری بیشتر برای مدل سازی روابط رقابتی در ارتباطات بی‌سیم به کمک پهپاد برای کنترل توان، تخصیص منابع، موقعیت‌یابی پهپادها و امنیت استفاده می‌شود. به عنوان مثال، دو پهپاد متعلق به اپراتورهای مختلف برای کسب و کار و پهپادهای نظامی با هم رقابت می‌کنند که سعی می‌کنند سیستم‌های ارتباطی دشمن را نظارت، پارازیت یا ضدجمع کنند.

مرجع[۳۶]، مساله موقعیت یابی پهپادها را به منظور به حداکثر رساندن پوشش دستگاه های تلفن همراه (به عنوان مثال، تعداد دستگاه‌های تلفن همراه متصل) مورد مطالعه قرار داد. در این حالت، تعدادی دستگاه‌های تلفن همراه به طور تصادفی روی زمین حرکت می‌کنند. سه پهپاد انتخاب می‌کنند که در سلول فعلی خود دایره بزنند یا به سمت مرکز حرکت کنند و یک سلول مجاور بر اساس تعداد دستگاه‌های تلفن همراهی که پشتیبانی می‌کند. اگر اقدام مربوطه انتخاب شود، ماتریس پرداخت حاوی مقادیر پوشش هر پهپاد است. سپس همه بازیکنان استراتژی های خود را به طور همزمان با یافتن یک تعادل نش از این ماتریس پرداخت انتخاب می‌کنند. پوشش پهپادها ۱۱,۹ درصد بهبود یافته است. همچنین نشان داده شده است که این طرح تئوری بازی در مقایسه با سه سناریوی پوشش تک پهپاد از نظر انرژی کارآمدتر است.

با این حال، تنها سه پهپاد در آن مقاله در نظر گرفته شده است و معروف است که بازی با فرم معمولی با افزایش تعداد بازیکنان، مشکل مقیاس پذیری دارد. علاوه بر این، هنگام مقایسه بازده توان، تنها توان ارتباطی در نظر گرفته می‌شود در حالی که توان مورد نیاز برای حرکت در نظر گرفته نمی‌شود.

نتیجه‌گیری و ارائه پیشنهادات

با توجه به پیشرفتِ انقلابی در فناوری اطلاعات ، تغییراتِ زیادی در ماهیتِ نبردها ایجاد شده است و همچنین رشدِ روز افزونِ فناوری اطلاعات و هوش مصنوعی در میدان‌های نبرد، منافع استراتژیکِ هر یک از طرفین در صحنه نبرد، بستگی زیادی به عملکردِ هوشمندانه طرف مقابل دارد. یکی از روش‌های نوین در این حوزه، که می‌تواند نتایجِ اثربخشی را در صحنه‌های نبردِ نوین داشته باشد، استفاده و به کارگیری روش‌های نظریه بازی‌ها می‌باشد. بنابراین در این پژوهش، به معرفی نظریه بازی‌ها پرداخته شد و ضمن تعریف مفاهیم پایه در این حوزه، دسته-بندی انواع بازی‌ها براساس معیارهای مختلف، بیان شد. سپس انواعِ نحوه‌ی نمایش بازی‌ها در محیط مساله معرفی گردید. همچنین چار چوب انواع بازی‌های دونفره، چند نفره، تکراری و

تکاملی، جمعیتی و بر ضد طبیعت، بیان شد تا بتوان برای به کارگیری هر یک از این انواع بازی‌ها، در مسایل مختلف، تصمیم‌گیری شود. در پایان، کاربردهای بازی‌ها در ارتباط پهپادها مبتنی بر شبکه‌های بی‌سیم بیان شد.

استفاده از روش بازی‌های تکاملی مبتنی بر بهینه‌سازی چندین هدفه، در به کارگیری در صحنه‌های نبرد می‌تواند به عنوان کارهای آتی پیشنهاد و پیاده‌سازی گردد.

قدردانی

از کلیه اساتید و خبرگانی که با صرف وقت ارزشمند خود در خصوص تکمیل و انجام این پژوهش ما را یاری نمودند کمال تشکر را داریم.

منابع

[1]. Juve K. The use of massive multiplayer online games to evaluate C4I systems. NAVAL POSTGRADUATE SCHOOL MONTEREY CA; 2004.
[2]. Halpern JY. Computer science and game theory: A brief survey. arXiv preprint cs/0703148. 2007.
[3]. Roy S, Ellis C, Shiva S, Dasgupta D, Shandilya V, Wu Q, editors. A survey of game theory as applied to network security. 2010 43rd Hawaii International Conference on System Sciences; 2010: IEEE.
[4]. Shi H-Y, Wang W-L, Kwok N-M, Chen S-Y. Game theory for wireless sensor networks: a survey. Sensors. 2012;12(7):9055-97.
[5]. Wang Y, Wang Y, Liu J, Huang Z, Xie P, editors. A survey of game theoretic methods for cyber security. 2016 IEEE First International Conference on Data Science in Cyberspace (DSC); 2016: IEEE.
[6]. Do CT, Tran NH, Hong C, Kamhoua CA, Kwiat KA, Blasch E, et al. Game theory for cyber security and privacy. ACM Computing Surveys (CSUR). 2017;50(2):1-37.
[7]. Yang G. Game theory-inspired evolutionary algorithm for global optimization. Algorithms. 2017;10(4):111.
[8]. Habib MA, Moh S. Game theory-based routing for wireless sensor networks: A comparative survey. Applied Sciences. 2019;9(14):2896.
[9]. Liu Z, Luong NC, Wang W, Niyato D, Wang P, Liang Y-C, et al. A survey on applications of game theory in blockchain. arXiv preprint arXiv:190210865. 2019.
[10]. Riahi S, Riahi A. Game theory for resource sharing in large distributed systems. International Journal of Electrical & Computer Engineering (2088-8708). 2019;9(2).
[11].Zargaryan M, Gevorgyan D. Distributed Algorithms and Game Theory.
[12]. Gintis H. Game theory evolving: Princeton university press; 2009.
[13]. Keen D, Andersson R. Double games: Success, failure and the relocation of risk in fighting terror, drugs and migration. Political Geography. 2018;67:100-10.

[14]. Vamvoudakis KG, Lewis FL. Online solution of nonlinear two-player zero-sum games using synchronous policy iteration. International Journal of Robust and Nonlinear Control. 2012;22(13):1460-83.
[15]. Bailey JP, Piliouras G. Multi-agent learning in network zero-sum games is a Hamiltonian system. arXiv preprint arXiv:190301720. 2019.
[16]. Jaśkiewicz A, Nowak AS. Non-zero-sum stochastic games. Handbook of dynamic game theory. 2016:1-64.
[17]. Ummels M. Stochastic multiplayer games: Theory and algorithms: Amsterdam University Press; 2010.
[18]. Gámez M, López I, Rodríguez C, Varga Z, Garay J. Game-theoretical model for marketing cooperative in fisheries. Applied Mathematics and Computation. 2018;329:325-38.
[19]. Zhang H, Lian J, Wang H, editors. Improve the Cooperative Level of Population via Individual Recognition Model. 2019 Chinese Control And Decision Conference (CCDC); 2019: IEEE.
[20]. Watson M, Bozgeyikli L, editors. Introduction to Game Theory via an Interactive Gameplay Experience. Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion; 2019.
[21]. Gokhale CS, Traulsen A. Evolutionary multiplayer games. Dynamic Games and Applications. 2014;4(4):468-88.
[22]. Newton J. Evolutionary game theory: A renaissance. Games. 2018;9(2):31.
[23]. Muell MR, Guillory WX, Kellerman A, Rubio AO, Scott-Elliston A, Morales O, et al. Gaming natural selection: Using board games as simulations to teach evolution. Evolution. 2020;74(3):681-5.
[24]. Cui J, Liu Y, Nallanathan A. Multi-agent reinforcement learning-based resource allocation for UAV networks. IEEE Transactions on Wireless Communications. 2019;19(2):729-43.
[25]. Messous M-A, Senouci S-M, Sedjelmaci H, Cherkaoui S. A game theory based efficient computation offloading in an UAV network. IEEE Transactions on Vehicular Technology. 2019;68(5):4964-74.
[26]. Yan S, Peng M, Cao X. A game theory approach for joint access selection and resource allocation in UAV assisted IoT communication networks. IEEE Internet of Things Journal. 2018;6(2):1663-74.
[27]. Shiri H, Park J, Bennis M, editors. Massive autonomous UAV path planning: A neural network based mean-field game theoretic approach. 2019 IEEE Global Communications Conference (GLOBECOM); 2019: IEEE.
[28]. Zhang K, Yang Z, Başar T. Multi-agent reinforcement learning: A selective overview of theories and algorithms. Handbook of Reinforcement Learning and Control. 2021:321-84.
[29]. Klaine PV, Nadas JP, Souza RD, Imran MA. Distributed drone base station positioning for emergency cellular networks using reinforcement learning. Cognitive computation. 2018;10(5):790-804.
[30]. Gholamnezhad P, Mazloum J. UAV optimal routing based on reference vector guided evolutionary algorithm. Journal of Aeronautical Engineering. 2021;23(1):44-55.
[31]. Kim H, Park J, Bennis M, Kim S-L, editors. Massive UAV-to-ground communication and its stable movement control: A mean-field approach. 2018 IEEE

19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC); 2018: IEEE.
[32]. Xu Y, Ren G, Chen J, Luo Y, Jia L, Liu X, et al. A one-leader multi-follower Bayesian-Stackelberg game for anti-jamming transmission in UAV communication networks. Ieee Access. 2018;6:21697-709.
[33]. Xiao L, Xie C, Min M, Zhuang W. User-centric view of unmanned aerial vehicle transmission against smart attacks. IEEE Transactions on Vehicular Technology. 2017;67(4):3420-30.
[34]. Choudhary G, Sharma V, You I, Yim K, Chen R, Cho J-H, editors. Intrusion detection systems for networked unmanned aerial vehicles: a survey. 2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC); 2018: IEEE.
[35]. Saad W, Han Z, Basar T, Debbah M, Hjorungnes A, editors. A selfish approach to coalition formation among unmanned air vehicles in wireless networks. 2009 International Conference on Game Theory for Networks; 2009: IEEE.
[36]. Charlesworth PB, editor Using non-cooperative games to coordinate communications UAVs. 2014 IEEE Globecom Workshops (GC Wkshps); 2014: IEEE.