نشریه علمی «علوم و فناوری‌های پدافند نوین»
سال سیزدهم، شماره ۱، بهار ۱۴۰۱؛ ص ۳۴- ۲۳

علمی- پژوهشی

الگوریتم‌هایی با پیچیدگی زمانی چندجمله‌ای برای حل مسائل بازی امنیتی مجموع صفر و مجموع ناصفر

سمانه اسماعیلی[۱]، حسن حسن‌پور[۲*]، حمید بیگدلی[۳]

۱.دانشجوی دکتری، ۲.دانشیار دانشگاه بیرجند ۳.استادیار دانشگاه فرماندهی و ستاد آجا

(دریافت: ۱۸/ ۱۴۰۰/۰۸، پذیرش: ۱۴۰۱/۰۲/۱۶)

چکیده

با توجه به اهمیت مسئله امنیت، تخصیص بهینه نیرو از موضوعات مورد توجه پژوهشگران است. در دو دهه گذشته، شاخه جدیدی از نظریه بازی به نام بازی امنیتی برای محاسبه سیاست دفاعی بهینه با موفقیت برای مسائل امنیتی به‌کار گرفته شده است. در این بازی ها علاوه بر محدودیت منابع، عکس‌العمل منطقی مهاجم به هر راهبرد مدافع نیز درنظر گرفته می شود. پیش از این با تحلیل نظریه بازی، مسائلی بهینه‌سازی به‌منظور تخصیص بهینه نیرو ارائه شده و الگوریتم‌هایی نیز پیشنهاد شده‌اند که برای هر نوع بازی امنیتی و در هر شرایطی کارایی ندارند. در این مقاله الگوریتمی با زمان اجرای چندجمله‌ای برای محاسبه میزان پوشش بهینه اهداف ارائه شده است. اساس کار الگوریتم، گسترش مجموعه اهدافی موسوم به مجموعه حمله است که در مجموعه بهترین پاسخ‌های مهاجم قرار می‌گیرند؛ و درنهایت محدود کردن این مجموعه به هدفی با بیشینه عایدی مدافع است. در ادامه، بازی امنیتی مجموع صفر معرفی شده است که در آن با انتخاب هر راهبرد مدافع، مجموع عایدی مدافع و مهاجم صفر است. ثابت می‌شود که برای محاسبه جواب بهینه در این بازی، کافی است بزرگ‌ترین مجموعه حمله محاسبه شود. بر این اساس، الگوریتمی زمان چندجمله‌ای برای این نوع بازی نیز ارائه شده است.

کلیدواژه‌ها: تخصیص بهینه نیرو، بازی امنیتی، بازی مجموع ناصفر، بازی مجموع صفر، الگوریتم زمان چندجمله‌ای.

Some Polynomial-time Algorithms for Solving Zero-Sum and Nonzero-Sum Security Game Problems

S. Esmaeeli, H. Hasanpour*, H. Bigdeli
University of Birjand
(Received: 09/11/2021; Accepted: 06 /05/2022)

Abstract

Given the importance of security, optimal force allocation is a topic of interest to researchers. Over the past two decades, a new branch of game theory called the security game, has been successfully used as a method to calculate the optimal defense policy for security issues. In these games, in addition to the limitations of security resources, the logical reaction of the attacker to any defender's strategy, is considered. Formerly, some optimization problems have been proposed to optimize the force allocation using the game theory, and some algorithms have been proposed that do not work for all types of security games and in all situations. In this paper, an algorithm with polynomial execution time is proposed to calculate the optimal coverage of the targets. The basis of this algorithm is expanding a set of targets called the attack set, which is included in the set of the attacker's best responses; and finally, limiting this set to a target with the maximum defender's payoff. Next, a zero-sum security game is introduced, in which by choosing any defender's strategy, the sum of defender's and attacker's payoffs are zero. It is proved that to calculate the optimal strategy in this game, it is enough to calculate the largest attack set. Accordingly, a polynomial-time algorithm is also proposed for this type of the game.

Keywords: Optimal Power Allocation, Security Game, Non-Zero Sum Game, Zero Sum Game, Polynomial Time Algorithm

*Corresponding Author E-mail: hhassanpour@birjand.ac.ir

۱. مقدمه

در اغلب مسائل امنیتی دنیای واقعی، محدودیت منابعی چون بودجه، پرسنل، تجهیزات و ... مانع از فراهم آوردن پوشش کامل اهداف مورد تهدید است. مسئله دیگر این است که برنامه‌ریزی قطعی تخصیص منابع به یک مهاجم فرضی اجازه می دهد تا راهبرد مدافع را مشاهده کند و احتمالاً از هر الگوی قابل پیش‌بینی در آن بهره ببرد تا یک حمله برنامه‌ریزی شده بهتر انجام دهد. در تحلیل نظریه بازی، تخصیص منابع در دفاع از اهداف بر اساس ارزش هدف برای مدافع و مهاجم و احتمال حمله به هدف از سوی مهاجم مورد ارزیابی قرار می‌گیرد. روش‌های گذشته بدون درنظر گرفتن رفتار منطقی مهاجمان، قادر به مدل‌سازی رفتار منطقی بازیکنان نیستند، در نتیجه کارایی راه‌حل برای مدافع کاهش می‌یابد. نظریه بازی با درنظر گرفتن ارزش هدف برای مدافع و مهاجم و رفتار منطقی مهاجم، فنون جذابی را برای یافتن یک راهبرد تخصیص منابع ارائه می‌دهد.

در دهه‌های اخیر، نظریه بازی در حل مسئله تأمین امنیت در مقیاس بزرگ مانند سواحل و بنادر [۱]، فرودگاه‌ها [۲ و ۳]، شبکه‌های جاده‌ای شهری [۴]، شبکه‌های حمل و نقل [۵ و ۶]، امنیت سایبری [۱۰-۷] و گشت‌زنی [۱۱ و ۱۲] بسیار مورد توجه بوده است. نقاط مهم این دستگاه‌ها معمولاً زیرساخت‌های آسیب‌پذیر بزرگی هستند که چنانچه مورد تهاجم قرار گیرند، عملکرد دستگاه با اختلال روبه‌رو می‌شود و باعث از دست دادن بودجه زیادی در هر ساعت می‌شود. ضمن اینکه حمله به این زیرساخت‌ها ممکن است تلفات جانی زیادی نیز در پی داشته باشد. این‌ها بخشی از دلایلی است که سرمایه‌گذاری برای تأمین امنیت در مناطق آسیب‌پذیر را ضروری می‌کند. سرمایه‌گذاری در منابع متفاوتی چون تجهیزات، پرسنل، گشت سگ، غربالگری مسافر، ایستگاه‌های بازرسی تردد و منابع دفاعی، بسته به اهداف نیازمند تأمین امنیت، انجام می‌شود. اما پوشش امنیتی کامل مناطق در معرض تهدید همیشه ممکن نیست، بنابراین یک سیاست تخصیص منابع کارآمد مورد نیاز است. نظریه بازی یک رویکرد ریاضی منطقی برای تخصیص منابع امنیتی برای به حداکثر رساندن تأثیر آن‌ها فراهم می‌کند. به‌عبارتی نظریه بازی ابزارهای محاسباتی برای پرداختن به مسئله امنیت را در اختیار ما قرار می‌دهد.

در یک حوزه امنیتی، مدافع باید دائماً از مجموعه‌ای از اهداف با استفاده از تعداد محدودی از منابع دفاع کند و به حملات احتمالی مهاجم پاسخ دقیق بدهد. اگر مدافع را در نقش رهبر و مهاجم را در نقش پیرو درنظر بگیریم، توصیف یک بازی استاکلبرگ را خواهیم داشت. این مدل به مدافع کمک می‌کند که حتی اگر نوع دشمن مشخص نباشد، اقدامات خود را بهینه کند.

در سال‌های اخیر علاقه زیادی به مطالعه بازی‌های امنیتی برای بهینه‌سازی دفاعی و کاهش اثرات حملات به دستگاه‌ها و بخش‌های امنیتی مختلف شامل زیرساخت‌های حیاتی، دستگاه‌های مالی، مراکز حفاظتی و امنیتی شهری وجود دارد. از نظریه بازی می‌توان به‌عنوان یک ابزار ریاضی برای به حداکثر رساندن کارایی منابع امنیتی محدود استفاده کرد. ارتباط بین نظریه بازی و امنیت طی چندین دهه گذشته مورد بررسی قرار گرفته است [۱۵-۱۳].

کونیتزر و ساندولم [۱۶] الگوریتمی را برای یافتن راه‌حل بهینه با حل تعدادی از مسائل برنامه‌ریزی خطی ارائه داده‌اند. پاروچوری و همکاران [۱۷] الگوریتمی ابتکاری به نام ASAP[1] ارائه داده‌اند. لچفورد و همکاران [۱۸ و ۱۹] محاسبه راهبردهای استاکلبرگ را در بازی‌های تصادفی بررسی کرده‌اند. بیگدلی و حسن‌پور [۲۰] به بررسی بازی‌های چند هدفی در محیط قطعی پرداخته و از برنامه‌ریزی آرمانی برای محاسبه راهبرد بهینه مدافع استفاده کرده‌اند. بیگدلی و همکاران [۲۱] علاوه بر ارائه یک روش حل بازی‌های امنیتی چند هدفی با عایدی‌های فازی، کاربردی از این مدل را در ایجاد امنیت در ایستگاه‌های مترو ارائه داده‌اند. ترجو و همکاران [۲۳] مدلی را برای بازی‌های امنیتی استاکلبرگ بر اساس بازی‌های مارکوف پیشنهاد کرده‌اند.

اکثر بازی‌های امنیتی از بازی استاکلبرگ استفاده می‌کنند. زیرا نیروهای امنیتی معمولاً برای تخصیص نیروهای خود به سیاست‌های امنیتی خاصی متعهد هستند. بنابراین مهاجمان می‌توانند با نظارت چینش نیروهای مدافع در اهداف مختلف، از هر گونه ضعف احتمالی مدافع استفاده کنند. استفاده از مفهوم تعادل استاکلبرگ در بازی‌های امنیتی مورد توجه زیادی قرار گرفته است [۲۲ و ۲۳]. اسماعیلی و همکاران [۲۴ و ۲۵] به حل مسئله بازی امنیتی با منابع فریبنده با استفاده از بازی استاکلبرگ در محیط فازی پرداخته‌اند. کارگروه تامبه و همکاران، طیف گسترده‌ای از برنامه‌های کاربردی بازی استاکلبرگ را برای امنیت در ایالات متحده ارائه کرده‌اند. در تخصیص منابع گشت‌زنی، مدل PROTECT[2] با استفاده از چارچوب بازی استاکلبرگ برای تقویت امنیت دریایی بندر بوستون طراحی و از آوریل ۲۰۱۱ مورد استفاده قرارگرفته‌است [۲۶]. این مدل در حال حاضر در بنادر بوستون، لانگ‌بیچ، نیویورک، لس‌آنجلس و چند بندر دیگر مورد استفاده قرار گرفته است [۲۷]. برای مطالعه بیشتر، کونتی [۲۲] برخی الگوریتم‌ها و پژوهش‌هایی را که در زمینه بازی‌های استاکلبرگ انجام شده، گردآوری کرده است.

در این مقاله بازی‌های امنیتی مجموع ناصفر و مجموع صفر مورد بررسی قرار گرفته و الگوریتمی برای حل آن‌ها ارائه می‌شود.

[1] Agent Security via Approximate Policies
[2] Port Resilience Operational Tactical Enforcement to Combat Terrorism

در بازی امنیتی مجموع صفر در هر حمله مهاجم، مقدار عایدی که یکی از دو بازیکن به‌دست می‌آورد با مقداری که دیگری از دست می‌دهد برابر است. به‌عبارتی مجموع عایدی حاصل از انتخاب راهبردهای بازیکنان صفر است. در بخش نهایی بازی امنیتی مجموع صفر به‌عنوان حالت خاصی از بازی امنیتی مجموع ناصفر و الگوریتم حل آن معرفی می‌گردد.

۲. روش تحقیق

در این بخش ابتدا بازی امنیتی مجموع ناصفر، عناصر این بازی و توابع عایدی بازیکنان معرفی می‌شود. سپس به کمک مسئله برنامه‌ریزی خطی دوسطحی یک نقطه تعادل برای بازی امنیتی مجموع ناصفر محاسبه می‌گردد. درنهایت الگوریتمی برای محاسبه تعادل در این بازی ارائه می‌گردد. در ادامه بازی امنیتی مجموع ناصفر به اختصار بازی امنیتی نامیده می‌شود.

۲-۱. بازی‌های امنیتی مجموع ناصفر

نیروهای امنیتی همواره باید آمادگی مقابله با مهاجم را داشته باشند. از این‌رو قبل از حمله مهاجم راهبرد امنیتی خود را انتخاب می‌کنند. بنابراین از بازی استاکلبرگ برای مدل‌سازی بازی امنیتی استفاده می‌شود که در آن ابتدا یک بازیکن به‌عنوان رهبر راهبردی را برمی‌گزیند و سپس بازیکن دوم به‌عنوان پیرو با مشاهده راهبرد رهبر، راهبردش را اتخاذ می‌کند. در این بازی فرض بر این است که رهبر همیشه متعهد است. تعهد، پایبندی وی به تصمیم اولیه‌اش را پس از اعلام تصمیم تضمین می‌کند. به بیان دیگر رهبر بعد از اعلام تصمیم اولیه‌اش اجازه ندارد بازگردد و تصمیم خود را تغییر دهد و باید به حرکت خود پایبند باشد.

فرض کنید $I = \{1,2,\dots,n\}$ و $J = \{1,2,\dots,m\}$ به‌ترتیب مجموعه راهبردهای محض رهبر و پیرو باشند. یک راهبرد آمیخته برای هر بازیکن یک توزیع احتمال روی مجموعه راهبردهای محض اوست. اگر مجموعه راهبردهای آمیخته رهبر و پیرو به‌ترتیب با X و Y نمایش داده شوند، در این صورت داریم:

$$X = \{x \in \mathbb{R}^n | \sum_{i=1}^{n} x_i = 1, x_i \geq 0, i = 1,\dots,n\}$$

$$Y = \{y \in \mathbb{R}^m | \sum_{j=1}^{m} y_j = 1, y_j \geq 0, j = 1,\dots,m\}.$$

راهبردهای محض حالت خاصی از راهبردهای آمیخته هستند. بنابراین در ادامه راهبرد آمیخته به‌اختصار راهبرد نامیده می‌شود.

راهبردی (مجموعه راهبردهایی) که عایدی یک بازیکن را در برابر راهبرد ثابت بازیکن دیگر بیشینه می‌کند، بهترین پاسخ (مجموعه بهترین پاسخ‌های) وی نام دارد. فرض کنید $x \in X$ و $y \in Y$ به‌ترتیب راهبردهای اتخاذ شده رهبر و پیرو و $R(x)$ و $\mathrm{U}(x,y)$ به‌ترتیب نشان‌دهنده بهترین پاسخ بازیکن پیرو به راهبرد x بازیکن رهبر و عایدی او از انتخاب راهبرد y در برابر راهبرد x رهبر باشد. در این‌صورت داریم:

$$R(x) = \max_{y \in Y} \mathrm{U}(x,y).$$

در بازی استاکلبرگ هدف تعیین راهبرد بهینه رهبر است. برای حصول این هدف، ابتدا بیشینه عایدی پیرو به ازای راهبردهای مختلف رهبر به‌دست آمده و سپس عایدی رهبر روی بهترین پاسخ‌های پیرو بهینه می‌شود. جفت راهبردی را که از این طریق محاسبه می‌شود، جواب تعادل استاکلبرگ می‌نامند.

تعریف ۱: فرض کنید x و y راهبردهای اتخاذ شده رهبر و پیرو و $U_1(x,y)$ و $U_2(x,y)$ به‌ترتیب عایدی‌های این بازیکنان باشند. جفت راهبرد (x^*,y^*) را جواب تعادل بازی استاکلبرگ گوییم هرگاه جواب مسئله زیر باشد:

$$U_1(x^*,y^*) = \max_{x \in X} U_1(x, R(x))$$

که در آن، $R(x)$ نشان دهنده بهترین پاسخ بازیکن پیرو به راهبرد x بازیکن رهبر است.

بازی امنیتی که رقابت بین مدافعان و مهاجمان است، نوع خاصی از بازی استاکلبرگ است. مدافع می‌تواند با هر نوع تخصیص نیرو به اهداف مورد نظرش حملات مهاجم را جهت‌دهی کند. به‌عبارتی برنامه‌ریزی مختلف مدافع تعیین کننده حملات مهاجمان به اهداف است. مهاجم استدلال می‌کند و هدفی را که منابع دفاعی کافی برای محافظت در اختیار داشته باشد، کمتر مورد حمله قرار می‌دهد. این به این معنا نیست که تمایل داشته باشد به اهداف با پوشش کمتر یا فاقد پوشش حمله کند بلکه عایدی‌اش از حمله به هر هدف را نیز درنظر می‌گیرد. بنابراین مدافع باید بتواند تعیین کند که برای هر استقرار نیرو چه مقدار عایدی نصیب خودش و رقیبش (مهاجم) می‌شود. فرض کنید مدافع پوشش $\mathrm{c_t}$ را به هدف t اختصاص دهد. اگر $\mathrm{c_t} = 1$، هدف t کاملاً محافظت شده است و افزودن پوشش بیشتر به این هدف تغییری در تمایل مهاجم برای حمله به آن ایجاد نمی‌کند. در ضمن منابع دفاعی مدافع محدود است و افزودن منابع دفاعی به این هدف باعث تضییع نیروی دفاعی است. در این مقاله فرض می‌شود مدافع حداکثر m منبع دفاعی در اختیار دارد. هر تخصیص پوشش به اهداف، یک راهبرد برای مدافع است. بنابراین راهبرد مدافع بردار $C = (c_1, c_2, \dots, c_n)$ با محدودیت‌های زیر است:

$$\forall t \in T \quad 0 \leq c_t \leq 1 \ , \qquad \sum_{t=1}^{n} c_t \leq m.$$

پس از استقرار نیروهای دفاعی در اهداف، مهاجم با مشاهده عملکرد مدافع تصمیم می‌گیرد که به کدام هدف یا اهداف حمله کند. او می‌تواند کل توان هجومی خود را روی یک هدف متمرکز

کند (فقط به یک هدف حمله کند) یا این توان را روی اهداف مختلف تقسیم کند. پس راهبرد مهاجم، برداری مانند $A = (a_1, a_2, \dots, a_n)$ تعریف می‌شود به‌طوری که

$$\forall t \in T \quad a_t \geq 0\,, \qquad \sum_{t=1}^{n} a_t = 1$$

که در آن، a_t، درصد نیروی مهاجم در حمله به هدف t است.

جفت راهبرد (C, A) را یک نمایه راهبرد بازی می‌نامند. بدیهی است یک مهاجم منطقی پس از مشاهده راهبرد C مدافع، به هدفی (یا اهدافی) حمله می‌کند که بیشترین عایدی را برایش داشته باشد. انتخاب راهبردی با بیشترین عایدی از سوی مهاجم را بهترین پاسخ او می‌نامند. در هر نمایه راهبرد (C, A)، A بهترین پاسخ مهاجم به راهبرد C مدافع است.

در این بازی امنیتی، عایدی دو بازیکن در دو حالت وجود پوشش محافظتی و عدم وجود پوشش برای هر هدف t به‌ترتیب با $U_d^c(t)$ و $U_d^u(t)$ برای مدافع و $U_a^c(t)$ و $U_a^u(t)$ برای مهاجم نمایش داده می‌شود که با ماتریسی مشابه جدول (۱) نمایش داده می‌شود. این ماتریس توسط خبرگان در اختیار قرار می‌گیرد.

جدول ۱. ماتریس عایدی مدافع و مهاجم

هدف t	$c_t = 0$	$c_t = 1$
مهاجم	$U_a^u(t)$	$U_a^c(t)$
مدافع	$U_d^a(t)$	$U_d^c(t)$

بدیهی است چنانچه هدفی توسط یکی از منابع مدافع محافظت شود، عایدی مدافع به‌ازای آن هدف افزایش و عایدی مهاجم کاهش می‌یابد. عایدی مدافع و مهاجم از پوشش c_t برای هدف t به‌ترتیب زیر محاسبه می‌شوند:

$$U_d(t, c_t) = c_t U_d^c(t) + (1 - c_t) U_d^u(t) \qquad (۱)$$

$$U_a(t, c_t) = c_t U_a^c(t) + (1 - c_t) U_a^u(t) \qquad (۲)$$

همچنین فرض می‌شود $U_d^c(t) > U_d^u(t)$ و $U_a^c(t) < U_a^u(t)$، که فرضی منطقی است. عایدی بازیکنان به‌ازای تنها یک هدف محاسبه نمی‌شود. تخصیص منابع پوششی به اهدافی که مورد حمله قرار نمی‌گیرند برای مدافع حائز اهمیت نیست. زیرا تخصیص این منابع به اهداف در معرض حمله، عایدی بیشتری (زیان کمتری) برای مدافع در پی دارد. عایدی کلی مدافع و مهاجم به ازای انتخاب نمایه راهبرد (C, A)، به‌ترتیب به‌صورت زیر محاسبه می‌شود:

$$U_d(C, A) = \sum_{t=1}^{n} a_t U_d(t, c_t) \qquad (۳)$$

$$U_a(C, A) = \sum_{t=1}^{n} a_t U_a(t, c_t) \qquad (۴)$$

پس از انتخاب راهبرد C مدافع، مهاجم بر اساس رابطه (۴) بررسی می‌کند که عایدیش در حمله به کدام اهداف بیشینه می‌شود. تمایل مهاجم برای حمله به اهداف، بستگی به میزان پوشش اهداف مختلف دارد. ممکن است حمله به یک هدف بدون پوشش دفاعی، عایدی زیادی برای مهاجم نداشته باشد، اما پوشش دفاعی سایر اهداف به حدی باشد که مهاجم ترجیح دهد به هدف مذکور (هدف فاقد پوشش) حمله کند. برای هر راهبرد C مدافع، مجموعه اهداف با بیشترین عایدی برای مهاجم را با $\Gamma(C)$ نمایش داده و آن را مجموعه حمله می‌نامیم. به‌عبارتی

$$\Gamma(C) = \{t: U_a(t, c_t) \geq U_a(\acute{t}, c_t) \;\; \forall \acute{t} \in T\} \qquad (۵)$$

هر بردار پوشش C مجموعه حمله مخصوص به خود را دارد. زیرا با تغییر بردار پوشش، عایدی مهاجم از اهداف تغییر کرده و در نتیجه مجموعه اهداف با بیشینه عایدی برای مهاجم تغییر خواهد کرد. در حقیقت $\Gamma(C)$ مجموعه اهدفی است که با انتخاب راهبرد C مدافع، حمله به آن‌ها بهترین پاسخ مهاجم است. به‌عنوان مثال چنانچه اهداف فاقد پوشش دفاعی باشند، یعنی $C = 0$، برای هر هدف t، $U_a(t, c_t) = U_a^u(t)$ و داریم:

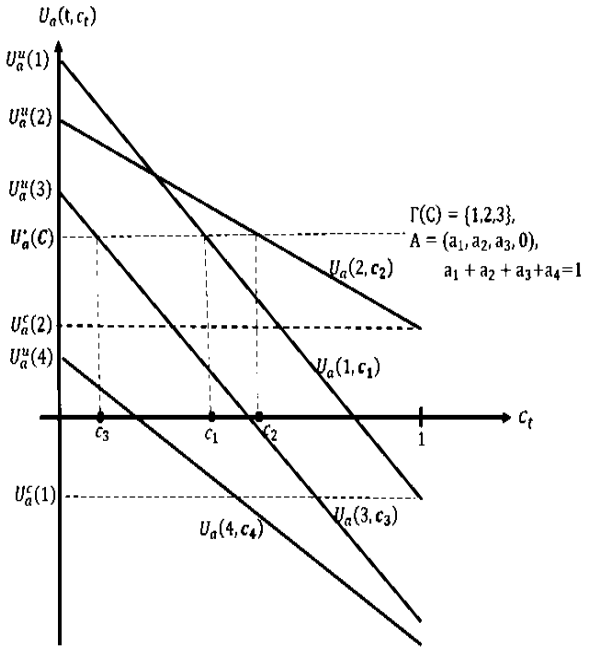
$$\Gamma(C) = \left\{ \acute{t}: U_a^u(\acute{t}) = \max_{t \in T} U_a^u(t) \right\} \qquad (۶)$$

بدیهی است که برای هر راهبرد C مدافع، مقادیر $U_a(t, c_t)$ برای هر $t \in \Gamma(C)$، یکسان است (هر چند مقادیر c_t برای این اهداف یکسان نباشد). این مقدار را با $U_a^*(C)$ نمایش می‌دهیم. بنابراین برای هر راهبرد C مدافع $U_a^*(C) = \max_{t \in T} U_a(t, c_t)$ و برای هر $t \in \Gamma(C)$، $U_a(t, c_t) = U_a^*(C)$.

برای هر $t \in T$، توابع عایدی معرفی شده در روابط (۱) و (۲)، توابع خطی وابسته به c_t هستند. تابع عایدی مهاجم را می‌توان به شکل زیر بازنویسی کرد:

$$U_a(t, c_t) = c_t (U_a^c(t) - U_a^u(t)) + U_a^u(t) \qquad (۷)$$

چون $U_a^c(t) < U_a^u(t)$، $U_a(t, c_t)$ به‌عنوان تابعی از c_t در بازه $[0,1]$ تابعی خطی با شیب $U_a^c(t) - U_a^u(t) < 0$ است و چنانچه $c_t \geq 1$، مقادیر کمتر از $U_a^c(t)$ را اختیار می‌کند. در این شرایط هدف t به‌طور کامل محافظت شده و بنابراین مهاجم علاقه‌ای ندارد به هدف t حمله کند. مجموعه حمله و مقدار $U_a^*(C)$ بر حسب راهبرد C مدافع، تغییر می‌کنند. شکل (۱)، نمودار عایدی مهاجم برای هر هدف را بر حسب c_t نشان می‌دهد. در این شکل فرض شده مجموعه اهداف $T = \{1,2,3,4\}$ بر حسب $U_a^u(t)$ به‌صورت نزولی مرتب شده‌اند. همچنین مجموعه $\Gamma(C)$، برای $C = (c_1, c_2, c_3, 0)$ مشخص شده است.



شکل ۱. نمودار عایدی مهاجم به ازای اهداف مختلف

مجموعه حمله به پوشش اهداف بستگی دارد. به‌عبارتی برای هر راهبرد C مدافع، مجموعه حمله متفاوت خواهد بود. به‌عنوان مثال در مجموعه حمله شکل (۱)، چنانچه میزان c_2 را افزایش دهیم، $U_a(2,c_2)$ کاهش یافته و بنابراین این هدف از مجموعه حمله خارج خواهد شد. در حقیقت $\Gamma(C)$ مجموعه بهترین پاسخ‌های مهاجم به راهبرد C مدافع است. در مجموعه حمله شکل (۱)، برای مهاجم تفاوتی نمی‌کند که به هدف ۱، ۲ یا ۳ حمله کند یا حمله خود را بین دو یا سه هدف از $\Gamma(C)$ تقسیم کند. بدیهی است که مهاجم به اهداف خارج از مجموعه حمله، حمله نخواهد کرد. به عبارتی برای $t \notin \Gamma(C)$ داریم $a_t = 0$. با توجه به روابط (۱) و (۲)، برای راهبرد C مدافع تنها اهداف مجموعه حمله در عایدی مدافع و مهاجم نقش دارند. زیرا با حمله به اهداف خارج از مجموعه حمله عایدی کمتر از $U_a^*(C)$ نصیب مهاجم می‌شود. همچنین مهاجم لزوماً به همه اهداف مجموعه حمله، حمله نخواهد کرد.

۲-۲. محاسبه یک نقطه تعادل برای بازی امنیتی مجموع ناصفر

هدف از حل بازی به‌دست آوردن نقطه تعادل است که در آن هیچ یک از بازیکنان با تغییر بازی خود نتوانند عایدی بیشتری به‌دست آورند. در بازی امنیتی به‌عنوان نوعی بازی استاکلبرگ، ابتدا مدافع چینش دفاعی خود را برمی‌گزیند و سپس مهاجم با توجه به میزان پوشش محافظتی هر هدف، راهبرد بهینه خود را اتخاذ می‌کند (به هدف یا اهدافی حمله می‌کند). برای به‌دست آوردن نقطه تعادل باید مدافع راهبردهای بهینه مهاجم به‌ازای هر راهبرد خود را در نظر بگیرد و سپس عایدی خود را بیشینه کند. با توجه به تعریف تعادل استاکلبرگ، تعادل بازی امنیتی به شکل زیر تعریف می‌شود:

تعریف ۲: فرض کنید $U_d(C,A)$ و $U_a(C,A)$ به‌ترتیب عایدی‌های مدافع و مهاجم به ازای جفت راهبرد (C,A) باشد. راهبرد (C^*,A^*) را جواب تعادل بازی امنیتی گوییم هرگاه جواب مسئله زیر باشد:

$$U_d(C^*,A^*) = \max_C U_d\left(C, \max_A U_a(C,A)\right) \qquad (۸)$$

از تعریف فوق و با توجه به محدودیت‌های راهبردهای مدافع و مهاجم، برای محاسبه یک نقطه تعادل بازی می‌توان مسئله برنامه‌ریزی دوسطحی (BLP) را حل کرد.

$$\text{(BLP)}: P_1: \max_C U_d(C,A) = \sum_{t=1}^{n} a_t U_d(t,c_t)$$
$$s.t. \quad \sum_{t=1}^{n} c_t \le m$$
$$0 \le c_t \le 1$$

که در آن، A جواب بهینه مسئله زیر است:

$$P_2: \max_A U_a(C,A) = \sum_{t=1}^{n} a_t U_a(t,c_t)$$
$$s.t. \quad \sum_{t=1}^{n} a_t = 1$$
$$a_t \ge 0$$

مدافع بررسی می‌کند که به ازای راهبردهای مختلف او، یعنی انتخاب هر بردار C شدنی مسئله P_1، مهاجم چه عکس‌العملی نشان می‌دهد. در مسئله P_1 مدافع راهبرد C را انتخاب می‌کند. مسئله P_2، به‌عنوان بخشی از محدودیت‌های مسئله P_1، جواب‌های شدنی را به بهترین پاسخ‌های مهاجم محدود می‌کند (بهترین پاسخ‌های مهاجم به‌ازای انتخاب C مدافع را به‌دست می‌دهد). سپس با درنظر گرفتن این پاسخ‌ها، راهبردی را اتخاذ می‌کند که بیشترین عایدی را برایش داشته باشد یعنی $U_d(C,A)$ را بیشینه کند.

با توجه به قیود مسئله P_2، برای مهاجم تفاوتی نمی‌کند که کل توان هجوم خود را روی یک هدف از مجموعه حمله متمرکز کند یا آن را روی اهداف مجموعه حمله تقسیم کند.

گزاره ۱: برای هر نمایه راهبرد (C,A) در جواب بهینه مسئله P_2، برای هر $t \notin \Gamma(C)$، $a_t = 0$.

اثبات: به برهان خلف فرض کنید اندیسی مانند $t_0 \in T$ موجود باشد که $t_0 \notin \Gamma(C)$ و $a_{t_0} \neq 0$. در این صورت چون $t_0 \notin \Gamma(C)$ داریم $U_a(t_0,c_{t_0}) < U_a^*(C)$. از طرفی به‌ازای هر $t \in \Gamma(C)$ داریم $U_a(t,c_t) = U_a^*(C)$. بنابراین با توجه به قیود P_2 به‌دست می‌آوریم:

$$
\begin{aligned}
U_a(C,A) &= \sum_{t\in T} a_t\, U_a(t,c_t)\\
&= \sum_{t\in \Gamma(C)} a_t\, U_a(t,c_t) + \sum_{\substack{t\notin \Gamma(C)\\ t\neq t_0}} a_t\, U_a(t,c_t)\\
&\quad + a_{t_0} U_a\left(t_0,c_{t_0}\right)\\
&< \sum_{t\in \Gamma(C)} a_t\, U_a^*(C) + \sum_{\substack{t\notin \Gamma(C)\\ t\neq t_0}} a_t\, U_a^*(C)\\
&\quad + a_{t_0} U_a^*(C)\\
&= U_a^*(C)\sum_{t\in \mathrm{T}} a_t = U_a^*(C)
\end{aligned}
$$

که با بهینگی (C,A) برای P_2 در تناقض است. بنابراین فرض خلف باطل و حکم ثابت شده است.

گزاره ۲: برای هر انتخاب راهبرد C مدافع، مقدار بهینه تابع هدف مسئله P_2، $U_a^*(C)$ است.

اثبات: تابع هدف مسئله P_2 به شکل زیر است:

$$U_a(C,A) = \sum_{t=1}^{n} a_t\, U_a(t,c_t)$$

طبق گزاره ۱، برای هر $t \notin \Gamma(C)$، $a_t = 0$. بنابراین در جواب بهینه مسئله P_2 داریم:

$$
\begin{aligned}
U_a(C,A) &= \sum_{t\in \Gamma(C)} a_t\, U_a(t,c_t) = \sum_{t\in \Gamma(C)} a_t\, U_a^*(C)\\
&= \sum_{t\in \mathrm{T}} a_t\, U_a^*(C) = U_a^*(C)
\end{aligned}
$$

آخرین تساوی از قیود مسئله P_2 نتیجه می‌شود.

از گزاره‌های فوق نتیجه می‌شود که مدافع عایدی بهینه را از بین اهدافی در مجموعه حمله به‌دست می‌آورد. زیرا بهترین پاسخ مهاجم انتخاب هدف یا اهدافی از مجموعه حمله است و برای هدفی که به آن حمله نشود $a_t = 0$. گزاره ۱ نتایج زیر را نیز به دنبال دارد.

نتیجه ۱: برای نمایه راهبرد (C,A)، اگر $t \in \mathrm{T}$ موجود باشد که $t \notin \Gamma(\mathrm{C})$ و $a_t \neq 0$، نمایه راهبرد (C,A)، یک جواب شدنی برای مسئله P_1 نیست.

اثبات: از گزاره ۱ نتیجه می‌شود که نمایه راهبردی که در آن $t \in \mathrm{T}$ موجود باشد که $t \notin \Gamma(\mathrm{C})$ و $a_t \neq 0$، جواب بهینه مسئله P_2 و در نتیجه جواب شدنی برای مسئله P_1 نیست.

نتیجه ۲: نمایه راهبرد (C,A) که در آن برای هر $t \in T$، $0 \le c_t \le 1$ و $\sum_{t=1}^{n} c_t \le m$، یک جواب شدنی برای مسئله P_1 است، اگر و تنها اگر $A \in \mathrm{A(C)}$ که

$$A(C) = \{(a_1,\dots,a_n): \textstyle\sum_{t\in \Gamma(C)} a_t = 1\,, a_t \ge 0\}$$

اثبات: فرض کنید (C,A) با شرایط داده شده، جوابی شدنی برای مسئله P_1 باشد. بنابراین باید برای مسئله P_2 نیز شدنی باشد و در نتیجه برای هر $t \in T$، داریم $a_t \ge 0$. همچنین از گزاره ۱، برای هر $t \notin \Gamma(C)$، $a_t = 0$. بنابراین $\sum_{t\in \Gamma(C)} a_t = 1$ که نتیجه می‌دهد $A \in \mathrm{A(C)}$.

برعکس فرض کنید $A \in \mathrm{A(C)}$. نمایه راهبرد (C,A) برای مسئله P_2 شدنی است. برای هر $\acute{t} \in \Gamma(C)$، $U_a(\acute{t},c_{\acute{t}}) = U_a^*(C)$ و برای هر $t \in \mathrm{T}$، $U_a(\acute{t},c_{\acute{t}}) \ge U_a(t,c)$. بنابراین نمایه راهبرد (C,A) که در آن $A \in \mathrm{A(C)}$ برای مسئله P_2 بهینه است با مقدار بهینه تابع هدف $U_a^*(C)$. بنابراین با در نظر گرفتن سایر فرضیات نتیجه ۲، نمایه راهبرد (C,A) برای مسئله P_1 یک جواب شدنی است.

نتیجه ۳: کاهش یا افزایش c_t برای هر $t \notin \Gamma(\mathrm{C})$، موجب افزایش عایدی مدافع نخواهد شد.

اثبات: طبق گزاره ۱، برای هر $t \notin \Gamma(C)$، $a_t = 0$، بنابراین با توجه به تابع هدف مسئله P_1، افزایش c_t در اهدافی که در مجموعه حمله نیستند، موجب افزایش عایدی مدافع نخواهد شد.

نتیجه ۴: اگر C راهبرد بهینه مدافع باشد در این صورت برای هر $t \notin \Gamma(\mathrm{C})$ داریم $c_t = 0$.

اثبات: بدیهی است.

با توجه به گزاره ۱، در جواب بهینه مسئله P_1 داریم:

$$U_d(C,A) = \sum_{t=1}^{n} a_t\, U_d(t,c_t) = \sum_{t\in \Gamma(C)} a_t U_d(t,c_t)$$

پس می‌توان با افزایش پوشش بعضی از اهداف مجموعه حمله، عایدی مدافع را افزایش داد. اما افزایش پوشش این اهداف تا جایی موجب افزایش عایدی مدافع می‌شود که هدف از مجموعه حمله خارج نشود یا هدفی با عایدی بیشتر برای مدافع وارد مجموعه حمله نشود. چون فقط اهداف مجموعه حمله در افزایش مقدار بهینه تابع هدف تأثیر دارند، مجموعه حمله تا حد ممکن گسترش داده می‌شود تا اهدافی با عایدی بیشتر برای مدافع وارد مجموعه حمله شوند. سپس در این مجموعه، پوشش هدفی با بیشترین عایدی برای مدافع به‌گونه‌ای تغییر داده می‌شود که فقط این هدف در مجموعه حمله باقی بماند.

۳. الگوریتم زمان چندجمله‌ای برای حل بازی امنیتی مجموع ناصفر

در این بخش الگوریتمی با زمان اجرای چندجمله‌ای برای تعیین نقطه تعادل بازی امنیتی مجموع ناصفر ارائه می‌شود. در روند الگوریتم، پوشش‌ها طوری به اهداف تخصیص داده می‌شوند که بزرگ‌ترین مجموعه حمله ممکن با پوشش‌های موجود ساخته

شود. سپس در این مجموعه، هدفی با بیشترین عایدی برای مدافع مشخص شده و پوشش این هدف به اندازه‌ای کاهش می‌یابد که مجموعه حمله به همین هدف محدود شود.

فرض کنید بزرگ‌ترین مجموعه حمله ممکن، به‌ازای راهبرد C مدافع ساخته شود. به‌ازای انتخاب راهبرد C، هدفی مانند t^* وجود دارد که $U_d(t^*, c_{t^*}) = \max_{t=1,\dots,n} U_d(t, c_t)$. بنابراین با توجه به اینکه مسئله مقید به قید $\sum_{t=1}^{n} a_t = 1$ است، تابع هدف مسئله P_1، ترکیب محدبی از $U_d(t, c_t)$ به‌ازای $t = 1, \dots, n$ بوده، $U_d(C, A)$ بیشینه مقدار خود را به‌ازای $a_{t^*} = 1$ اختیار می‌کند و داریم

$$Max\, U_d(C, A) = U_d(t^*, c_{t^*}).$$

اکنون کافی است مجموعه بهترین پاسخ‌های مهاجم به راهبردهایی مانند $A = (a_1, \dots, a_n)$ محدود شود که در آن $a_{t^*} = 1$ و برای $t \neq t^*$ ، $a_t = 0$. برای این کار با قرار دادن $c_{t^*} - \varepsilon$ که $\varepsilon > 0$، به جای c_{t^*}، تنها این هدف در مجموعه حمله باقی می‌ماند. با این روش $U_d(C, A)$ بیشینه مقدار خود را به‌دست خواهد آورد. ε توسط خبرگان و با توجه به نوع منابع پوششی و جداول عایدی بازیکنان تعیین می‌گردد. انتخاب ε باید به‌گونه‌ای باشد که علاوه بر تأثیرگذاری بر تمایل مهاجم، عایدی مدافع را کاهش چشمگیری ندهد. بنابراین باید

$$\forall\, t \in \Gamma(C) \quad U_d(t^*, c_{t^*} - \varepsilon) > U_d(t, c_t)$$

از رابطه (۱) باید برای هر $t \in \Gamma(C)$ داشته باشیم:

$$0 < \varepsilon < \left| \frac{U_d(t^*, c_{t^*}) - U_d(t, c_t)}{U_d^c(t^*) - U_d^u(t^*)} \right|.$$

در قضیه ۱ اثبات می‌شود که هیچ تغییری در بردار C به‌دست آمده از این روش، باعث افزایش عایدی مدافع نخواهد شد. بنابراین نمایه راهبرد (C, A) که از این روش محاسبه شود، یک نقطه تعادل است.

در روند فوق‌الذکر تا جایی که ممکن است، اهدافی که می‌توانند در عایدی مدافع (تابع هدف مسئله P_1) تأثیرگذار باشند وارد مجموعه حمله می‌شوند. محاسبه بزرگ‌ترین مجموعه حمله برای مسئله P_1، با پوشش‌دهی اهدافی که بهترین پاسخ مهاجم هستند، انجام می‌شود. فرض کنید اهداف بر حسب $U_a^u(t)$، به‌صورت نزولی مرتب شده باشند. ابتدا پوشش همه اهداف را صفر در نظر می‌گیریم. در حالتی که همه اهداف فاقد پوشش هستند، $\Gamma(C) = \{1\}$. سپس هدف ۱ با بیشینه $U_a^u(t)$ را تا جایی پوشش می‌دهیم که $U_a^u(2) = U(1, c_1)$. در این شرایط $\Gamma(C) = \{1,2\}$. توجه شود که افزایش بیشتر پوشش هدف ۱ موجب خروج این هدف از مجموعه حمله می‌شود. سپس پوشش اهداف ۱ و ۲ را تا جایی افزایش می‌دهیم که $U_a^u(3) = U(1, c_1) = U(2, c_2)$. این روند را تا جایی ادامه می‌دهیم که $\sum_{t\in\Gamma(C)} c_t = m$. مجموعه حمله محاسبه شده به این طریق بزرگ‌ترین مجموعه حمله ممکن برای مهاجم خواهد بود. زیرا برای ورود هدف جدیدی به مجموعه حمله، باید هدف مورد نظر بهترین پاسخ مهاجم برای حمله باشد. بدین منظور باید پوشش اهداف مجموعه حمله را افزایش دهیم تا مهاجم برای حمله به هدف مذکور، تمایل داشته باشد. اما از آنجا که $\sum_{t\in\Gamma(C)} c_t = m$، این کار امکان‌پذیر نیست. همچنین کاهش پوشش هر یک از اهدافی که وارد مجموعه حمله شده‌اند، باعث خروج سایر اهداف از مجموعه حمله و کوچک‌تر شدن آن می‌شود. به‌عنوان مثال در مجموعه حمله مشخص شده در شکل (۱)، چنانچه $c_1 + c_2 + c_3 = m$، این مجموعه بزرگ‌ترین مجموعه حمله ممکن است.

اگر به‌ازای راهبرد C مدافع که $\sum_{t\in\Gamma(C)} c_t = m$، مقدار $U_a^*(C)$ در دست باشد، می‌توان بزرگ‌ترین مجموعه حمله را مشخص و مقادیر پوشش بهینه را برای مدافع محاسبه کرد. زیرا محل برخورد خط افقی $U_a(t, c_t) = U_a^*(C)$ و نمودار تابع $U_a(t, c_t)$ برای هر $t \in T$ مقادیر c_t را به‌دست می‌دهد و اگر $t \notin \Gamma(C)$ آنگاه $U_a^u(t) < U_a^*(C)$ و در نتیجه

$$c_t = a_t = U_a(t, c_t) = 0.$$

درنهایت پس از محاسبه بزرگ‌ترین مجموعه حمله و مشخص شدن میزان پوشش اهداف در این مجموعه، t^* مشخص شده و نقطه تعادل و نیز مقدار بهینه تابع هدف مدافع به شیوه‌ای که توضیح داده شد، محاسبه می‌گردد.

بدون آنکه از کلیت مسئله کاسته شود، فرض کنید مجموعه T بر حسب $U_a^u(t)$ به‌صورت نزولی مرتب شده باشد و $\Gamma = \{1, \dots, p\}$ که $p < n$. برای سادگی در نوشتار فرض کنید $U_a^*(C) = y$. داریم:

$$(u_a^c(1) - u_a^u(1))c_1 + u_a^u(1) = y$$
$$\vdots$$
$$(u_a^c(p) - u_a^u(p))c_P + u_a^u(p) = y$$

بنابراین:

$$c_t = \frac{y - U_a^u(t)}{U_a^c(t) - U_a^u(t)} \quad \mathrm{t} = 1, \dots, \mathrm{p} \qquad (۹)$$

از طرفی $\sum_{t\in\Gamma(C)} c_t = m$ نتیجه می‌دهد که:

$$\sum_{t=1}^{p} \frac{y - U_a^u(t)}{U_a^c(t) - U_a^u(t)} = m \qquad (۱۰)$$

در رابطه (۱۰) تنها مجهول y است و به‌راحتی قابل محاسبه است. پس از محاسبه y از رابطه (۱۰)، می‌توان مقادیر c_t را برای $t = 1, \dots, p$ از روابط (۹) محاسبه کرد. برای $t \notin \Gamma(C)$ نیز $c_t = 0$.

برای محاسبه C به روش فوق نیاز داریم ابتدا مجموعه حمله را محاسبه کنیم. در الگوریتم (۱)، ابتدا مجموعه حمله مشخص شده و سپس مقدار y محاسبه می‌گردد.

الگوریتم ۱. محاسبه نقطه تعادل بازی امنیتی مجموع ناصفر

۱- ورودی: مجموعه T که بر حسب $U_a^u(t)$ به‌صورت نزولی مرتب شده و مقادیر $U_a^c(t)$، $U_d^u(t)$ و $U_d^c(t)$ بر اساس ترتیب نزولی $U_a^u(t)$.

۲- خروجی: بردار C، بردار حمله $A=(a_1,\dots,a_n)$ که در آن $a_{t^*}=1$ و برای $t\neq t^*$ ، $a_t=0$ و $U_d(t^*,c_{t^*})$ به‌عنوان بیشینه مقدار تابع هدف.

۳- قرار دهید $next=2$ و $\Gamma=\{1,2\}$.

۴- قرار دهید $y=U_a^u(next)$.

۵- اگر $t\in\Gamma$ وجود دارد که $U_a^c(t)<y$ قرار دهید $c_t=1$ و $m-1\longrightarrow m$ و $\Gamma\backslash\{t\}\rightarrow\Gamma$.

۶- با y مرحله ۳، مقادیر c_t را از روابط (۹) محاسبه کنید.

- اگر $\sum_{t\in\Gamma(C)}c_t<m$، قرار دهید

$\Gamma\cup\{t\}\rightarrow\Gamma$، $next+1\longrightarrow next$ و به گام ۴ بروید.

- اگر $\sum_{t\in\Gamma(C)}c_t>m$ قرار دهید

$\Gamma\backslash\{next\}\rightarrow\Gamma$ و به گام ۷ بروید.

- اگر $\sum_{t\in\Gamma(C)}c_t=m$ یا $\Gamma=\mathrm{T}$ به گام ۷ بروید.

۷- برای هر $t\notin\Gamma$ قرار دهید $c_t=0$.

۸- با Γ به‌دست آمده در گام ۶، مقدار y را از رابطه (۱۰) محاسبه کنید.

۹- با y به‌دست آمده در گام ۸، مقادیر c_t را برای هر $t\in\Gamma$ از رابطه (۹) محاسبه کنید.

۱۰- با C محاسبه شده، برای هر یک از اهداف مجموعه حمله، $U_d(t,c_t)$ را محاسبه کنید. t^* را هدفی با بیشینه $U_d(t,c_t)$ تعریف کرده و مقدار ε را در بازه زیر تعیین کنید:

$$0<\varepsilon<\min_{t\in\Gamma}\{\left|\frac{U_d(t^*,c_{t^*})-U_d(t,c_t)}{U_d^c(t)-U_d^u(t)}\right|\}$$

۱۱- میزان پوشش t^* را مساوی $c_{t^*}-\varepsilon$ قرار دهید.

قضیه ۱: نمایه (C,A)، محاسبه شده در الگوریتم (۱) یک نقطه تعادل است.

اثبات: با توجه به نتیجه ۲، (C,A) یک جواب شدنی برای مسئله P_1 است. کافی است نشان دهیم این جواب برای این مسئله بهینه است. نشان می‌دهیم هیچ تغییری در C موجب افزایش عایدی مدافع نمی‌شود. در روند الگوریتم اهداف به سه مجموعه افراز می‌شوند. نشان می‌دهیم تغییرات پوشش در هیچ یک از این اهداف موجب بهبود تابع هدف مسئله P_1 نمی‌شود.

۱- مجموعه اهدافی که طی الگوریتم وارد مجموعه Γ نمی‌شوند: طبق نتیجه ۳، کاهش یا افزایش c_t برای هر $t\notin\Gamma(C)$، موجب افزایش عایدی مدافع نمی‌گردد.

۲- مجموعه اهدافی که وارد Γ می‌شوند به جز هدف t^*: کاهش پوشش این اهداف موجب حضور آن‌ها در مجموعه حمله خواهد شد و چون برای اهدافی که وارد مجموعه Γ شده‌اند، $U_d(t^*,c_{t^*})=\max_C U_d(t,c_t)$، حضور این اهداف در تابع هدف با ضریب مثبت a_t، موجب افزایش تابع هدف نخواهد شد (مسئله مقید به $\sum_{t=1}^n a_t=1$ است). افزایش پوشش این اهداف نیز تأثیری بر تابع هدف نخواهد داشت، زیرا این اهداف در آخرین مرحله از مجموعه حمله خارج شده‌اند.

۳- هدف t^*: چون $U_d(t^*,c_{t^*})=\max_{t=1,\dots,n}U_d(t,c_t)$، کاهش c_{t^*} موجب افزایش مقدار بهینه تابع هدف نمی‌شود و افزایش پوشش c_{t^*}، موجب خروج آن از مجموعه حمله یا ورود هدف دیگری به Γ خواهد شد که این تغییر نیز کمکی به افزایش مقدار تابع هدف نمی‌کند.

مثال ۱: بازی امنیتی با دو منبع امنیتی مدافع، چهار هدف T={1,2,3,4} و جدول عایدی زیر را درنظر بگیرید (جدول ۲):

جدول ۲. ماتریس عایدی مدافع و مهاجم در مثال ۱

بازیکنان / اهداف	عایدی	مدافع	عایدی	مهاجم
	$c_t=1$	$c_t=0$	$c_t=1$	$c_t=0$
۱	۴	-۳	۰	۵
۲	۳	-۴	-۲	۴
۳	۶	-۵	-۱	۷
۴	۱	-۲	-۱	۱

با اجرای الگوریتم (۱) با داده‌های مسئله، بردار پوشش اهداف پس از اتمام گام ۸ به‌صورت زیر به‌دست می‌آید:

$$C=(0.77\,,0.57\,,0.73\,,0).$$

در این انتخاب راهبرد، عایدی مهاجم در حمله به اهداف ۱، ۲ و ۳ مساوی $U_a^*(C)=\mathrm{y}=1.15$ و عایدی مدافع در حمله به هر یک از این اهداف به ترتیب مساوی 2.39، -0.01 و 3.041 است. بنابراین $t^*=3$. در گام ۱۰ مقدار ε در بازه $\varepsilon\in(0,0.1)$ از تصمیم گیرنده درخواست می‌شود:

$$0<\varepsilon<\min\{|0.1|,|0.27|\}\Rightarrow 0<\varepsilon<0.1\ .$$

بنابراین بردار حمله مهاجم $A=(0,0,1,0)$ و بردار پوشش نهایی به‌صورت $C=(0/77\,,0/57\,,0/73-\varepsilon\,,0)$ است.

۴. محاسبه نقطه تعادل در مسئله بازی امنیتی مجموع صفر

در این بخش بازی امنیتی مجموع صفر معرفی می‌گردد، سپس به کمک ویژگی خاص این بازی‌ها الگوریتمی برای محاسبه پوشش بهینه ارائه می‌شود.

در بعضی از حملات، مهاجمان قصد تصاحب اهداف را دارند. اهدافی که مدافع از دست بدهد مهاجم به‌دست می‌آورد و چنانچه مدافع با حفاظت هوشمندانه، هدفی را برای خود حفظ کند، مهاجم با حمله به آن هدف نمی‌تواند آن را تصرف کند. این رقابت مثالی از یک بازی امنیتی مجموع صفر است. در این نوع بازی عایدی که یک بازیکن به‌دست می‌آورد، مساوی عایدی است که بازیکن دیگر از دست می‌دهد. به بیان ریاضی، فرض کنید C و A به ترتیب راهبردهای اتخاذ شده مدافع و مهاجم در یک بازی مجموع صفر باشند. در این صورت برای هر $t \in T$، $U_a(t, c_t) + U_d(t, c_t) = 0$. وجود این تساوی در بازی‌های مجموع صفر، حل این بازی را ساده‌تر می‌کند. در ادامه نشان می‌دهیم که برای محاسبه راهبرد بهینه مدافع در این بازی‌ها کافی است بزرگ‌ترین مجموعه حمله ممکن محاسبه شود.

گزاره ۳: در بازی امنیتی مجموع صفر، راهبرد C مدافع عایدی او را بیشینه می‌کند اگر و تنها اگر به‌ازای هر راهبرد $\acute{C}$ مدافع داشته باشیم $U_a^*(C) \le U_a^*(\acute{C})$.

اثبات: رابطه $U_a(t, c_t) = -U_d(t, c_t)$ برای بازی امنیتی مجموع صفر نتیجه می‌دهد که شرط لازم و کافی برای اینکه راهبرد C مدافع، عایدی مدافع را بیشینه کند این است که راهبرد C عایدی مهاجم را کمینه کند. بنابراین راهبرد C مدافع، عایدی مدافع را بیشینه می‌کند اگر و تنها اگر به ازای هر راهبرد $\acute{C}$ مدافع داشته باشیم $U_a^*(C) \le U_a^*(\acute{C})$.

قضیه ۲: در بازی امنیتی مجموع صفر، راهبرد C مدافع عایدی او را بیشینه می‌کند اگر و تنها اگر دو شرط زیر را داشته باشد:

۱- $\sum_{t\in\Gamma(C)} c_t = m$

۲- هر هدف $t \in T$ که $U_a^*(C) \le U_a^u(t)$، در مجموعه $\Gamma(C)$ حضور داشته باشد.

اثبات: فرض کنید راهبرد C عایدی مدافع را بیشینه کند. نشان می‌دهیم C باید دو شرط مذکور را داشته باشد. شرط $\sum_{t\in\Gamma(C)} c_t = m$ یک شرط لازم برای بهینه بودن راهبرد C است. زیرا انتخاب راهبرد $\acute{C}$ که $\sum_{t\in\Gamma(C)} \acute{c}_t < m$ باعث می‌شود مدافع بعضی منابع دفاعی را در اختیار داشته و بتواند با افزودن این منابع دفاعی روی اهداف مجموعه حمله، عایدی مهاجم را کاهش دهد. اکنون فرض کنید برای راهبرد $\acute{C}$ مدافع که $\sum_{t\in\Gamma(C)} \acute{c}_t = m$ هدف $\acute{t} \in T$ چنان موجود باشد که $U_a^*(\acute{C}) \le U_a^u(\acute{t})$ و $\acute{t} \notin \Gamma(\acute{C})$. اگر $\acute{c}_{\acute{t}} = 0$ آن‌گاه از رابطه (۴) داریم:

$$U_a^u(\acute{t}) = U_a(\acute{t}, c_{\acute{t}}) > U_a^*(\acute{C})$$

که با تعریف $U_a^*(\acute{C})$ در تناقض است. زیرا برای $\acute{t} \notin \Gamma(\acute{C})$ باید $U_a(\acute{t}, c_{\acute{t}}) \le U_a^*(\acute{C})$ و اگر $\acute{c}_{\acute{t}} > 0$، از آنجا که $\acute{t} \notin \Gamma(\acute{C})$، $a_{\acute{t}} = 0$، که نشان می‌دهد مدافع با انتخاب این راهبرد، بخشی از منابع را هدر داده است. زیرا منابع را به هدفی اختصاص داده که مهاجم به آن حمله نمی‌کند. بنابراین راهبرد C برای مدافع بهینه نیست.

برعکس، فرض کنید راهبرد C بهینه نباشد. نشان می‌دهیم هیچ تغییری در بردار C باعث افزایش عایدی مهاجم نمی‌شود. کاهش پوشش هر هدف در C موجب خروج بعضی اهداف از مجموعه حمله و افزایش $U_a^*(C)$ می‌شود که طبق گزاره ۳، باعث کاهش عایدی مدافع می‌گردد. همچنین از آنجا که $\sum_{t\in\Gamma(C)} c_t = m$، افزایش پوشش اهداف این مجموعه نیز امکان‌پذیر نیست.

نتیجه ۵: اگر راهبرد C مدافع عایدی او را بهینه کند آنگاه بزرگ‌ترین مجموعه حمله ممکن را برای مهاجم می‌سازد.

اثبات: اگر راهبرد C مدافع عایدی او را بیشینه کند برای ورود هدفی جدید مانند $\acute{t} \in T$ به مجموعه حمله مهاجم، باید هدف مورد نظر، بهترین پاسخ مهاجم برای حمله باشد. با افزایش پوشش هدف $\acute{t}$ نمی‌توان این هدف را به مجموعه حمله افزود. زیرا طبق شرط دوم قضیه ۲ باید $U_a^*(C) > U_a^u(\acute{t})$ (در غیر این صورت $\acute{t}$ در مجموعه حمله حضور داشته است). بنابراین باید پوشش اهداف مجموعه حمله را افزایش دهیم تا مهاجم برای حمله به هدف $\acute{t}$ تمایل داشته باشد. اما از آنجا که $\sum_{t\in\Gamma(C)} c_t = m$، این امر امکان‌پذیر نیست.

در قضیه ۲ نشان داده شد که راهبرد C مدافع که $\sum_{t\in\Gamma(C)} c_t = m$ برای مدافع بهینه است به شرط آن که هر هدف مانند $t \in T$ که $U_a^*(C) \le U_a^u(t)$، در این مجموعه حضور داشته باشد. راهبرد مذکور بزرگ‌ترین مجموعه حمله ممکن را برای مهاجم می‌سازد. بنابراین برای محاسبه راهبرد بهینه مدافع باید بزرگ‌ترین مجموعه حمله ممکن برای مهاجم که شرایط قضیه ۲ را داشته باشد، ساخته شود.

به روش الگوریتم (۱)، بزرگ‌ترین مجموعه حمله ممکن، راهبرد C متناظر با آن و مقدار $U_a^*(C)$ محاسبه می‌گردد. مقدار بهینه عایدی برای مدافع نیز $U_d(C, A) = -U_a^*(C)$ است. در الگوریتم (۱) و پس از محاسبه مقادیر c_t در گام ۹، موارد ذکر شده قابل محاسبه هستند. بنابراین برای محاسبه نقطه تعادل بازی امنیتی مجموع صفر گام ۱ تا ۹ الگوریتم (۱) اجرا می‌شود. خروجی الگوریتم بردار C محاسبه شده در گام ۹ و بردار

$A = (a_1, \dots, a_n)$ است که در آن برای هر $t \in \Gamma$ ، $0 \le a_t \le 1$ و برای هر $t \notin \Gamma$ ، $a_t = 0$. همچنین مقدار بهینه عایدی مدافع $U_d(C, A) = -\mathrm{y} = -U_a^*(C)$ است.

الگوریتم ۲. الگوریتم محاسبه نقطه تعادل در بازی امنیتی مجموع صفر

۱-ورودی: مجموعه T که بر حسب $U_a^u(t)$ به‌صورت نزولی مرتب شده و مقادیر $U_a^c(t)$، بر اساس ترتیب نزولی $U_a^u(t)$.
۲-خروجی: بردار C و $U_d(t^*, c_{t^*})$ به‌عنوان بیشینه عایدی مهاجم ($-U_d(t^*, c_{t^*})$ به‌عنوان بیشینه عایدی مدافع) و بردار حمله $A = (a_1, \dots, a_n)$ که در آن $\sum_{t\in\Gamma} a_t = 1$ و برای هر $t' \notin \Gamma$، $a_{t'} = 0$.
۳-گام‌های ۳ تا ۹ الگوریتم (۱) را تکرار کن.

در الگوریتم (۲)، بردار حمله به‌طور دقیق مشخص نمی‌شود. احتمال حمله مهاجم به اهداف خارج از مجموعه حمله صفر و احتمال حمله به هر یک از اهداف مجموعه حمله یکسان است.

مثال ۲: بازی امنیتی مجموع صفر با دو منبع امنیتی و جدول عایدی زیر را در نظر بگیرید (جدول برای مهاجم نوشته شده است، جدول ۳):

جدول ۳. ماتریس عایدی مهاجم در مثال ۲

بازیکن / اهداف	عایدی مهاجم $c_t = 1$	$c_t = 0$
۱	-۳	۵
۲	-۳	۶
۳	-۴	۴
۴	-۴	۵

با اجرای الگوریتم (۲)، مقادیر پوشش اهداف به‌صورت زیر محاسبه می‌شود:

$$C = (0/53, 0/59, 0/41, 0/47)$$

همچنین $U_a^*(C) = 0/735$ به‌دست می‌آید. $\Gamma = T$ و بنابراین مهاجم ممکن است به هر یک از اهداف حمله کند و عایدی مدافع درصورت حمله به هر یک از اهداف $U_d(A, C) = -0/735$ است.

۵. نتیجه‌گیری

از مهم‌ترین مسائل پیش روی مدافعان امنیت، محدودیت نیرو و تخصیص بهینه آن به اهداف نیازمند محافظت است. روش‌های بهینه‌سازی در پژوهش‌های گذشته فقط منافع مدافعان را درنظر گرفته و نیروهای موجود را به اهداف مختلف تخصیص داده‌اند. در حالی که بازی امنیتی یک رقابت دوطرفه بین مهاجمان و مدافعان است. مهاجمانی که قادر به تفکر و تصمیم‌گیری عقلایی هستند و در برابر راهبردهای مختلف مدافع، پاسخ منطقی می دهند. نظریه بازی‌ها در چند دهه اخیر به‌عنوان روشی برای مدل‌سازی رقابت‌ها بسیار مورد توجه قرار گرفته است. بازی بین مدافع و مهاجم که به بازی امنیتی شهرت یافته است، به محاسبه راهبرد بهینه برای مدافع در شرایطی می‌پردازد که منابع محدود است و امکان پوشش همه اهداف برای مدافع وجود ندارد. پیش از این با تحلیل نظریه بازی، مسائل برنامه‌ریزی جهت بهینه‌سازی تخصیص بهینه نیرو ارائه شده است. الگوریتم‌هایی نیز پیشنهاد شده‌اند که برای هر نوع بازی امنیتی و در هر شرایطی کارایی ندارند. در این مقاله الگوریتمی طراحی شده است که بازی امنیتی را در زمان چندجمله‌ای حل می‌کند. اساس الگوریتم بر این اصل استوار است که مهاجم به اهدافی حمله می‌کند که بیشترین عایدی را برایش دارند. این اهداف به‌ترتیب مقدار عایدی وارد مجموعه‌ای به‌نام مجموعه حمله می‌شوند. گسترش مجموعه حمله تا آنجا ادامه دارد که با توزیع پوشش‌های محافظتی روی این اهداف، پوشش‌های امنیتی مدافع به اتمام برسد یا تمام اهداف وارد مجموعه حمله شده باشند. در ادامه روند الگوریتم، هدفی که عایدی مدافع را بیشینه کند، در مجموعه حمله باقی می‌ماند. در بخش نهایی، بازی امنیتی مجموع صفر معرفی گردیده و الگوریتمی با زمان اجرای چندجمله‌ای برای محاسبه راهبرد بهینه برای مدافع ارائه شد. در بازی امنیتی مجموع صفر در هر حمله مهاجم، مقدار عایدی که یکی از دو بازیکن به‌دست می‌آورد (سود یک بازیکن) مساوی مقداری است که دیگری از دست می دهد (زیان بازیکن دیگر). ثابت شد راهبردی برای مدافع بهینه است که به‌ازای انتخاب هر راهبرد مدافع، کمترین عایدی را برای مهاجم داشته باشد و فقط بزرگ‌ترین مجموعه حمله ممکن برای مهاجم این شرط را داراست. الگوریتم ارائه شده بزرگ‌ترین مجموعه حمله ممکن را محاسبه کرده و با توجه به میزان منابع دفاعی موجود، پوشش بهینه را برای هر یک از اهداف محاسبه می‌کند.

۶. مراجع

[1] An, B.; Shieh, E.; Tambe, M.; Yang, R.; Baldwin, C.; DiRenzo, J.; Maule, B.; Meyer, G. "PROTECT: A Deployed Game Theoretic System for Strategic Security Allocation for the United States Coast Guard"; Ai Magazine 2012, 96-96.

[2] Hunt, K.; Agarwal, P.; Zhuang, J. "Technology Adoption for Airport Security: Modeling Public Disclosure and Secrecy in an Attacker-Defender Game"; Reliab. Eng. Syst. Safe 2021, 107355.

[3] Pita, J.; Tambe, M.; Kiekintveld, C.; Cullen, S.; Steigerwald, E. "Guards: Game Theoretic Security Allocation on a National Scale"; Tenth Int. Conf. on Autonomous Agents and Multiagent Syst. 2011, 37-44.

[4] Jain, M.; Korzhyk, D.; Vaněk, O.; Conitzer, V.; Pěchouček, M.; Tambe, M. "A Double Oracle Algorithm for Zero-sum Security Games on Graphs"; Tenth Int. Conf. on Autonomous Agents and Multiagent Syst. 2011, 327-334.

[5] Gnecco, G.; Hadas, Y.; Sanguineti, M. “Some Properties of Transportation Network Cooperative Games”; Networks 2019, 74, 161-173.

[6] Takalloo, M. “Game Theory Approaches for Transportation Problems”; PhD Diss., University of South Florida, 2020.

[7] Bansal, G.; Sikdar B. “Security Service Pricing Model for UAV Swarms: A Stackelberg Game Approach”; IEEE Conf. Comput. 2021, 1-6.

[8] Οικονομάκης, Π. “Strategic Military Deception Prerequisites of Success in Technological Environment”; 2015.

[9] Zhang, Y.; Malacaria, P. “Bayesian Stackelberg Games for Cyber-Security Decision Support”; Decis. Support Syst. 2021, 113599.

[10] Yuan, Y.; Sun, F.; Liu, H. “Resilient Control of Cyber-physical Systems against Intelligent Attacker: a Hierarchal Stackelberg Game Approach”; Int. J. Syst. Sci. 2016, 47, 2067-2077.

[11] Basilico N.; Celli A.; De Nittis G.; Gatti N. “Coordinating Multiple Defensive Resources in Patrolling GGames with Alarm Systems”; 16th Conf. Autonomous Agents and MultiAgent Syst. 2017, 678-686.

[12] Garrec, T. “Continuous Patrolling and Hiding Games”; Eur. J. Oper. Res. 2019, 277, 42-51.

[13] Albarran, S. E.; Clempner, J. B. “A Stackelberg Security Markov Game Based on Partial Information for Strategic Decision Making Against Unexpected Attacks”; Eng. Appl. Artif. Intel. 2019, 81, 408-19.

[14] Anwar, F.; Khan, B. U. I.; Olanrewaju, R. F.; Pampori, B. R.; Mir, R. N. “A Comprehensive insight into Game Theory in Relevance to Cyber Security”; Indonesian J. Elect. Eng. Informatics 2020, 8, 189-203.

[15] Sinha, A.; Fang, F.; An, B.; Kiekintveld, C.; Tambe, M. “Stackelberg Security Games: Looking beyond a Decade of Success”; Proc. Twenty-Seventh International Joint Conference on Artificial Intelligence, 2018.

[16] Conitzer, V.; Sandholm, T. “Computing the Optimal Strategy to Commit to”; Seventh ACM Conf. Elect. Commerce 2006, 82-90.

[17] Paruchuri, P.; Pearce, J. P.; Tambe, M.; Ordonez, F.; Kraus, S. “An Efficient Heuristic Approach for Security Against Multiple Adversaries”; 6th Conf. Autonomous Agents and Multiagent Syst., 2007, 1-8.

[18] Letchford, J.; MacDermed, L.; Conitzer, V.; Parr, R.; Isbell, C. L. “Computing Optimal Strategies to Commit to in Stochastic Games”; Twenty-Sixth AAAI Conf. Artif. Intel., 2012.

[19] Letchford, J.; Vorobeychik, Y. “Computing Randomized Security Strategies in Networked Domains”; Twenty-Fifth AAAI Conf. Artif. Intel. 2011.

[20] Bigdeli, H.; Hassanpour, H.; Tayyebi, J. “Optimistic and Pessimistic Solutions of Single and Multi-Objective Matrix Games with Fuzzy Payoffs and Analysis of Some Military Cases”; Adv. Defence Sci. & Technol. 2017, 8, 133-145 (In Persian).

[21] Bigdeli, H.; Hassanpour, H.; Tayyebi, J. “Multiobjective Security Game with Fuzzy Payoffs”; Iran J. Fuzzy Syst. 2019, 16, 89-101.

[22] Conti, S. “Algorithms for Finding Leader-Follower Equilibrium with Multiple Followers”; PhD Thesis, Politecnico Di Milano, 2013.

[23] Trejo, K. K.; Clempner, J. B.; Poznyak, A. S. “A Stackelberg Security Game with Random Strategies Based on the Extraproximal Theoretic Approach”; Eng. Appl. Artif. Intel. 2015, 37, 145-153.

[24] Esmaeeli, S.; Hassanpour, H.; Bigdeli, H. “Deception in Multi Attacker Security Game with Nonfuzzy and Fuzzy Payoffs”; Iranian Journal of Numerical Analysis and Optimization 2022, 12, 542–566.

[25] Esmaieli, S.; Hassanpour, H.; Bigdeli, H. “Lexicographic Programming for Solving Security Game with Fuzzy Payoffs and Computing Optimal Deception Strategy”; Defensive Future Study Researches J. 2020, 5, 89-108 (In Persian).

[26] Tambe, M.; Jiang, A. X.; An, B.; Jain, M. “Computational Game Theory for Security: Progress and Challenges”; AAAI Spring Symposium on Applied Computational Game Theory, 2014.

[27] An, B.; Ordóñez, F.; Tambe, M.; Shieh, E.; Yang, R.; Baldwin, C.; DiRenzo III, J.; Moretti, K.; Maule, B. Meyer, G. “A Deployed Quantal Response-based Patrol Planning System for the US Coast Guard”; Interfaces 2013, 43, 400-20.