Scientific Journal of Electronical & Cyber Defence
Vol. 11, No. 1, Spring 2023, Serial No. 41

Distributed Solving of Weapon Target Assignment Problem using Learning Automata

M. Shokoohi

* Instructor, Department of Computer Engineering, Technical and Vocational University, Tehran, Iran

(Received: 03/05/2021, Accepted: 01/01/2022)

ABSTRACT

This article presents a method to solve the weapon target assignment problem, which is one of the problems of distributed constraint optimization. The previous methods do not guarantee the convergence problem properly and when faced with scale increase, they do not work correctly and effectively. Also, some of these methods solve the weapon target assignment problem in a centralized manner. While the method presented in this article solves the problem in a decentralized and distributed manner with better "accuracy" and "speed". The present article solves the weapon target assignment problem by using learning automata, which is a relatively simple method and requires little information, and the results of implementations show that when the scale of the problem becomes larger, the proposed method it solves the problem with a better speed than other methods, so that the objective function is minimized. Also, this method can well address the shortcomings of previous methods without the need for other exploratory methods in real-time multi-agent environments.

Keywords: Weapons target assignment problem, Distributed problems, Learning Automata, Multi-agent Systems

* Corresponding Author Email: M-shokoohi@tvu.ac.ir

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Publisher: Imam Hussein University © Authors

نشریه علمی "پدافند الکترونیکی و سایبری"
سال یازدهم، شماره ۱، بهار ۱۴۰۲، ص ۴۷-۵۵

علمی – پژوهشی

ارائه یک راهکار توزیع‌شده برای مسئله تخصیص جنگ‌افزار به هدف با استفاده از آتوماتای یادگیر

مریم شکوهی
مربی، گروه مهندسی کامپیوتر، دانشگاه فنی و حرفه ای، تهران، ایران
(دریافت: ۱۴۰۰/۱۱/۰۱، پذیرش: ۱۴۰۱/۰۵/۲۴)

چکیده

در این مقاله روشی برای حل مسئلهٔ تخصیص جنگ‌افزار به هدف که یکی از مسائل بهینه‌سازی محدودیت توزیع‌شده است، ارائه می‌گردد. روش‌های قبل مشکل همگرایی را به‌درستی تضمین نمی‌کنند و زمانی که با افزایش مقیاس مواجه می‌شوند درست و مؤثر کار نمی‌کنند. همچنین برخی از این روش‌ها به‌صورت متمرکز مسئله تخصیص جنگ‌افزار به هدف را حل می‌کنند. درحالی‌که روش ارائه شده در این مقاله مسئله را به‌صورت غیرمتمرکز و توزیع شده با "دقت" و "سرعت" بهتری حل می‌کند. مقاله حاضر با استفاده از آتوماتای یادگیر که روشی نسبتاً ساده است و به اطلاعات کمی نیاز دارد، به حل مسئلهٔ تخصیص جنگ‌افزار به هدف می‌پردازد و نتایج پیاده‌سازی‌ها نشان می‌دهد، زمانی که مقیاس مسئله بزرگ‌تر می‌شود، روش پیشنهادی با سرعت بهتری نسبت به روش‌های دیگر، مسئله را حل کرده، به‌طوری‌که تابع هدف کمینه می‌شود. همچنین این روش به‌خوبی می‌تواند کمبودهای روش‌های قبلی را بدون نیاز به روش اکتشافی دیگر در محیط‌های چندعامله بلادرنگ پاسخگو باشد.

کلیدواژه‌ها: مسئله تخصیص جنگ‌افزار به هدف، مسئله توزیع‌شده، آتوماتای یادگیر، سیستم‌های چندعامله

۱- مقدمه

باتوجه‌به نقش و اهمیت کلیدی آرایش پویای تجهیزات و سایت‌های دفاعی در صحنه نبرد و استفاده‌ی بهینه از تجهیزات نظامی، تعیین تخصیص مناسب و بهینه جنگ‌افزارها به اهداف[1] (دشمن) از اهمیت بالایی برخوردار است [۱]. مسئلهٔ WTA ذاتا یک مسئله NP[2] کامل است [۲]. یعنی برای حل آن هیچ الگوریتم شناخته شده قابل اجرا در زمان چندجمله‌ای وجود ندارد. بنابراین الگوریتم‌های زیادی برای حل این مسئله ارائه شده است. برخی الگوریتم‌هایی که برای حل مسئلهٔ WTA به کار می‌روند، ذاتا متمرکز[3] هستند، یعنی یک سیستم مرکزی در مورد تخصیص جنگ‌افزارها به اهداف، تصمیم می‌گیرد. اما در کاربردهای واقعی، لازم است باتوجه‌به پراکندگی اهداف در محیط و افزایش تعداد آنها، هر یک از جنگ‌افزارها به‌صورت توزیع‌شده[4] و غیرمتمرکز تصمیم به تخصیص اهداف داشته باشند.

کارهای زیادی در زمینه مسئلهٔ تخصیص جنگ‌افزارها به اهداف انجام شده است؛ تا قبل از دهه ۸۰، الگوریتم‌ها به الگوریتم شمارش ضمنی[5]، شاخه و الگوریتم[6] و برنامه‌نویسی پویا[7] محدود شده است. این الگوریتم‌ها ساده هستند، اما در صورتی که مقیاس مسئله بزرگ شود، زمان اجرا به‌صورت نمایی افزایش می‌یابد [۳-۶]. با گسترش تکنولوژی برخی الگوریتم‌های بهینه مانند شبکه-های عصبی[8]، الگوریتم ژنتیک[9]، جستجوی ممنوعه[10] و الگوریتم شبیه‌سازی تبرید[11]، از دهه ۸۰ در قرن ۲۰ توسعه یافتند و برخی راه‌حل‌های جدید برای مسائل پیچیده ارائه شدند. الگوریتم‌های شبکه‌عصبی و الگوریتم ژنتیک بهبودیافته در مسئله WTA به کارگرفته شدند. با این حال، گاهی اوقات این الگوریتم‌ها ناپایدار هستند و همیشه همگرا نمی‌شوند. از این رو، راه‌حل ممکن است به‌صورت محلی بهینه باشد اما به‌صورت سراسری بهینه نشود [۷-۱۰].

با وجود روش‌های مختلف برای حل WTA، کارهای انجام شده دچار مجموعه‌ای از چالش‌ها و ضعف‌ها هستند که استفاده مطلوب از آنها را با مشکل مواجه می‌کنند. در بسیاری از روش‌ها، هماهنگی میان جنگ‌افزارها را در نظر نگرفته‌اند و فقط بر روی انتساب یک جنگ‌افزار متمرکز شده‌اند [۱۱، ۱۲]. برخی دیگر از الگوریتم‌ها مسائل را در مقیاس کوچک حل کرده‌اند، اما مسائل در مقیاس بزرگ به‌خوبی حل و فصل نشده است. در کارهای انجام شده، مسئله به‌صورت متمرکز در نظر گرفته شده است [۷، ۱۲]، درحالی‌که باتوجه‌به محیط عملیاتی و لزوم کارکرد بهتر و موثرتر جنگ‌افزارها، لازم است هریک از جنگ‌افزارها به‌صورت

* رایانامه نویسنده مسئول: M-shokoohi@tvu.ac.ir

[1] Weapon Target Assignment (WTA)
[2] Non-Deterministic Polynomial
[3] centralize
[4] distributed
[5] Implicit Enumeration
[6] Branch and Bound
[7] Dynamic Programming
[8] Neural Network
[9] Genetic Algorithm
[10] Tabu Search
[11] Simulated Annealing

* این مقاله یک مقاله با دسترسی آزاد است که تحت شرایط و ضوابط مجوز Creative Commons Attribution (CC BY) توزیع شده است.
© نویسندگان
ناشر: دانشگاه جامع امام حسین (ع)

توزیع‌شده مسئله تخصیص را برعهده گیرند [۱۱-۱۳]. در [۱۴] به تخصیص منابع در جنگ الکترونیک پرداخته شده است.

با وجود چنین چالش‌هایی در مسئله WTA، ارائه روشی جدید که توانایی بهبود و رفع موارد مذکور را داشته باشد، ضرورت دارد. همچنین در این مقاله، توجه به لزوم هماهنگی میان عامل‌ها در مدل ارائه شده و بررسی یک مدل غیرمتمرکز برای حل این مسئله، بررسی خواهد شد.

مسئله تخصیص جنگ‌افزار به هدف، شامل اختصاص بهینه جنگ‌افزار به هدف‌ها (دشمن) است به صورتی که مقدار باقی-مانده هدف‌ها، بعد از عملیات حداقل شود [۳]. در این کار فرض بر این است که جنگ‌افزارهایی از انواع مختلف وظیفه مقابله با اهداف مهاجم را برعهده دارند. جنگ‌افزارها به‌صورت غیرمتمرکز برای این وظیفه گماشته شده‌اند. در این رویکرد، موقعیت جنگ-افزارها و اهداف مهاجم، فاصله جنگ‌افزارها از اهداف، ارزش هر هدف مهاجم، احتمال انهدام اهداف توسط جنگ‌افزارها، حضور و یا عدم حضور اهداف مهاجم در محیط در نظر گرفته شده است. باتوجه‌به محدودیت در تعداد جنگ‌افزارها و ارزش اهداف و موقعیت آنها، می‌خواهیم تخصیص جنگ‌افزارها به اهداف را به‌صورت بهینه انجام دهیم تا میزان انهدام اهداف افزایش یابد.

در این مقاله برای حل مسئله تخصیص به‌صورت غیرمتمرکز و توزیع‌شده از آتوماتای یادگیر بهره گرفته شده است، به‌طوری‌که برای هر جنگ‌افزار در یک مسئله تخصیص، یک آتوماتای یادگیر در نظر گرفته شده است. ویژگی‌های همگرایی بلندمدت و پویایی یادگیری این آتومات‌ها به‌خوبی موردمطالعه قرار گرفته‌اند [۱۵، ۱۶]. علاوه بر این، درحالی‌که آتوماتای یادگیر یادگیرندگان نسبتاً ساده‌ای هستند، از لحاظ ارتباط بین آتوماتاها انعطاف پذیری زیادی ایجاد می‌کنند و امکان طراحی عوامل متشکل از یک مجموعه بزرگ آتوماتاهای به هم پیوسته را می‌دهند. این مزایا بیشتر با این واقعیت تکمیل می‌شوند که الگوریتم‌های آتوماتا به اطلاعات کمتری نیاز دارند و حتی نشان داده شده است که در محیط‌هایی با قابلیت مشاهده جزئی نیز کار می‌کنند [۱۷، ۱۸]، به همین دلیل آنها را برای سیستم‌های چندعامله مناسب می-کند.

ساختار مقاله به این شرح است: در بخش ۲ و ۳ پیشینه و مقدمه‌ای کوتاه از آتوماتای یادگیر، مسئله تخصیص جنگ‌افزار به هدف و کارهای انجام شده در این زمینه‌ها بیان می‌شود. در بخش ۴ و ۵ روش پیشنهادی و الگوریتم ارائه شده بر اساس آتوماتای یادگیر بازگو خواهد شد، به‌طوری‌که تنظیمات یادگیری و فرضیات استفاده شده، در این بخش نشان داده می‌شود. پیاده-سازی الگوریتم و نتایج اجرای آن در بخش ۶ بیان خواهد شد. در پایان مقاله در بخش ۷ نتیجه‌گیری و بیان کارهای آینده خواهد بود.

۲- پیشینه پژوهش

تحقیقات زیادی دررابطه‌با مسئله WTA وجود دارد. تقریباً تمامی آنها یک مسئله مشابه را تعریف کرده‌اند، اما تفاوت‌هایی در رویکرد راه‌حل‌ها وجود دارد. این راه‌حل‌ها معمولاً بر اساس یک‌سری متغیرها، از جمله قابلیت اهداف و جنگ‌افزارها، استراتژی دفاع، شرایط جنگی کنونی، دفاع لایه‌ای، حفاظت از سرمایه‌های خودی و یا انهدام اهداف دشمن و یا هر دو، ارزیابی تهدید باتوجه‌به قابلیت‌ها و تصمیم‌ها می‌باشند. همه این روش‌ها WTAرا به‌عنوان یک مسئله انتساب سخت[1] معرفی می‌کنند و می‌توان یک مثال از فرمول‌بندی جامع ریاضی برای مسئله را در [۳] پیدا کرد.

آهوجا و همکاران [۱۹] و لی [۲۰] با استفاده از روش فراابتکاری «جستجوی همسایگی در مقیاس بسیار بزرگ» راه‌حل‌های امکان‌پذیر آگاهانه را بهبود می‌بخشند. این راه‌حل یک جستجوی اکتشافی برای یافتن مؤثر یک راه‌حل باکیفیت است.

روزنبرگر و همکاران مسئله تخصیص را با استفاده از الگوریتم شاخه‌وحد[2] حل کردند [۲۱]. کلاین نیز با استفاده از این الگوریتم توانست راه‌حل‌های بهینه را برای مسئله تخصیص جنگ‌افزار به هدف بیابد [۲۲]. بوگدانوویچ در مقاله‌ی [۲۳] الگوریتمی توسعه داده و از آن برای جستجوی راه‌حل بهینه استفاده می‌کند.

توککوز و بولکان [۲۴] نتایج الگوریتم‌های ژنتیک، شبیه‌سازی تبرید، جستجوی همسایگی متغیر[3]، و جستجوی تابو را مقایسه کردند. جوهانسون و فالکمن [۲۵] از بهینه‌سازی ازدحام ذرات[4] استفاده می‌کنند و نتایج محاسباتی را با الگوریتم‌های ژنتیک، MMR و جستجوی کامل[5] مقایسه می‌کنند. برخی دیگر تأثیر حسگرها بر احتمال شناسایی موشک‌های ورودی را در نظر گرفته‌اند [۲۶, ۲۷].

جهت حل متمرکز مسئله تخصیص جنگ‌افزار به هدف، از الگوریتم‌های هوشمند [۲۸، ۲۹]، نظریه بازی [۳۰، ۳۱]، نمودار تصمیم بیزین [۳۲]، با الگوریتم فراابتکاری مبتنی برGRASP [۳۳] و روش‌های دیگر برای بهینه‌سازی سرعت محاسبه و انتساب هدف منطقی استفاده می‌شود، اما انعطاف‌پذیری سیستم با روش متمرکز مناسب نیست؛ برخی از نویسندگان (به‌عنوان مثال [۳۴]) به سادگی عامل‌ها را به‌عنوان برنامه‌ریزان متمرکزی در نظر گرفته‌اند که برنامه‌های خود را به عامل‌های دیگر منتشر می‌کنند، پس از آن حراج[6] و یا سازوکارهای مذاکره[7]، مسائل باقی را حل می‌کنند. اما اگر ارتباطات بین عامل‌ها، محدود، هزینه‌بر و

[1] Hard Assignment
[2] Branch and bound
[3] Variable Neighborhood Search
[4] Particle Swarm Optimization
[5] exhaustive search
[6] auction
[7] Negotiation

بدون امنیت باشند، کارایی این روش ضعیف خواهد بود.

در مقالات [۳۵، ۳۶] از الگوریتم توزیع‌شده برای بهینه‌سازی ارتباطات و محاسبات سیستم استفاده می‌شود، اما مشکل آن عدم آگاهی از وضعیت کلی سیستم است. این دو سازوکار تا حدی بهینه‌سازی مکانیزم تخصیص هدف را انجام می‌دهند. پیچیدگی و تضاد منابع مبارزه، تهدید اهداف، محیط ناپایدار و موارد دیگر به ندرت در [۳۷، ۳۸] در نظر گرفته شده است.

به‌طورکلی تحقیقات انجام شده بر روی مسئله WTA تا به امروز، یا بر روی حل مسئله WTA در حالت‌های خاص تمرکز داشته‌اند و یا بر روی توسعهٔ روش‌های اکتشافی مسئله WTA کار می‌کنند. علاوه بر این، ازآنجاکه هیچ الگوریتم دقیقی برای حل مسئلهٔ انتساب جنگ‌افزار به هدف در دسترس نیست، دقت راه‌حل‌های به‌دست‌آمده توسط این الگوریتم‌های اکتشافی مشخص نیست.

این بخش به بررسی و تعریف اجمالی مدل آتوماتای یادگیر و مسئلهٔ تخصیص جنگ‌افزار به هدف می‌پردازد.

۲-۱- آتوماتای یادگیر

آتومات‌های یادگیر[1] مدل‌های انتزاعی هستند که در محیطی تصادفی عمل نموده و قادر هستند که بر اساس ورودی‌های دریافت شده از محیط، احتمال انجام عملیات خود را به‌روز درآورده تا بتوانند از این طریق کارایی خود را بهبود بخشند [۳۹].

محیط را می‌توان توسط سه‌تایی $E \equiv \{\alpha, \beta, c\}$ تعریف نمود که $\alpha = \{\alpha_1, \alpha_2, .., \alpha_r\}$ مجموعه ورودی‌ها و $\beta = \{\beta_1, \beta_2, .., \beta_r\}$ مجموعه خروجی‌ها و $c = \{c_1, c_2, .., c_r\}$ مجموعه احتمال‌های جریمه شدن می‌باشند. هرگاه β_i دومقداری باشد، محیط از نوع p می‌باشد. در چنین محیطی $\beta_i = 1$ به‌عنوان جریمه و $\beta_i = 0$ به‌عنوان پاداش در نظر گرفته می‌شود. c_i احتمال این است که عمل α_i نتیجه نامطلوب داشته باشد.

آتوماتای یادگیر با ساختار متغیر توسط 4تائی $\{\alpha, \beta, p, T\}$ نشان داده می‌شود که در آن $\alpha = \{\alpha_1, \alpha_2, .., \alpha_r\}$ مجموعه عمل‌های آتوماتا، $\beta = \{\beta_1, \beta_2, .., \beta_r\}$ مجموعه ورودی‌های آتوماتا، $p = \{p_1, p_2, .., p_r\}$ بردار احتمال انتخاب هر یک از عمل‌ها و $p(n+1) = T[\alpha(n), \beta(n), p(n)]$ الگوریتم یادگیری می‌باشد. اگر $\beta = \{0,1\}$ باشد، محیط P-model خواهد بود.

الف -پاسخ مطلوب

$$p_g(\tau+1) = p_g(\tau) + a\left[1 - p_g(\tau)\right]$$
$$p_q(\tau+1) = (1-a)p_q(\tau) \qquad (۱)$$
$$\forall q \quad q \neq g$$

ب -پاسخ نامطلوب

$$p_g(\tau+1) = (1-b)p_g(\tau)$$
$$p_q(\tau+1) = {}^{b}/_{r-1} + (1-b)p_q(\tau) \qquad (۲)$$
$$\forall q \quad q \neq g$$

روابط (۱) و (۲) الگوریتم‌های یادگیری خطی در آتوماتای با ساختار متغیر با پاداش دومقداری است. در این نوع از آتوماتاها، اگر عمل α_i در مرحله nام انتخاب شود و پاسخ مطلوب از محیط دریافت نماید، احتمال $p_i(n)$ افزایش یافته و سایر احتمال‌ها کاهش یافته و برای پاسخ نامطلوب احتمال $p_i(n)$کاهش یافته و سایر احتمال‌ها افزایش می‌یابند. در هر حال، تغییرات به گونه‌ای صورت می‌گیرد تا حاصل جمع $p_i(n)$-ها همواره مساوی یک باقی بماند.

در روابط فوق، r تعداد اقدام‌های آتوماتای یادگیر j، τ زمان انجام عمل i، a پارامتر پاداش و b پارامتر جریمه می‌باشد. باتوجه‌به مقادیر a و b سه حالت زیر را می‌توان در نظر گرفت. زمانی که a و b با هم برابر باشند، الگوریتم را L_{RP}[2] می‌نامیم. وقتی که b از a خیلی کوچکتر باشد، الگوریتم را $L_{R\varepsilon P}$[3] می‌نامیم وقتی که b مساوی صفر باشد، الگوریتم را L_{RI}[4] می‌نامیم. حافظه و زمان مورد نیاز برای پیاده‌سازی آتوماتاهای یادگیر با ساختار ثابت $O(1)$ و برای آتوماتاهای یادگیر با ساختار متغیر $O(m)$ می‌باشد که m تعداد اعمال آتوماتاست [۳۹].

۲-۲- تخصیص جنگ‌افزار به هدف

مسئله تخصیص جنگ‌افزار به هدف مبتنی بر هدف[5]، انتساب بهینهٔ M جنگ‌افزار به N هدف است به‌طوری‌که آسیب مورد انتظار اهداف، حداکثر شود [۴۰، ۴۱]. فرض کنید M جنگ‌افزار و N هدف مستقل وجود دارد. متغیر تصمیم‌گیری x_{ij} به این صورت تعریف می‌شود که:

$$x_{ij} = \begin{cases} 1: & \text{if target } j \text{ assigned to weapon } i \\ 0: & \text{else} \end{cases} \qquad (۳)$$
$$i = 1,2,\dots,M \quad , \quad j = 1,2,\dots,N$$

باتوجه‌به اینکه جنگ‌افزار i درگیر هدف j است، یعنی $x_{ij} = 1$، بنابراین نتیجه تصادفی است و یک توزیع برنولی فرض می‌شود:

$$P(\text{target } j \text{ is destroyed by weapon } i) = P_{ij}$$
$$P(\text{target } j \text{ is not destroyed by weapon } i) = 1 - P_{ij}$$

فرض کنید که هر تعامل جنگ‌افزار، مستقل از هر تعامل دیگر است. بنابراین نتایج تعامل‌ها یک توزیع برنولی مستقل است. به هر هدف یک عدد حقیقی مثبت V_j اختصاص داده می‌شود که اولویت بین اهداف را نشان می‌دهد. فرض کنید که اهداف می‌توانند به کلاس‌هایی تقسیم شوند و هر کلاس دارای یک مقدار

[1] Learning Automata

[2] linear reward-penalty
[3] linear reward-ε-penalty
[4] linear reward-inaction
[5] Target based Weapon Target Assignment

منحصر به فرد برای تصمیم‌گیرنده است. مجموعه‌ای از همهٔ مقادیر کلاس در $\mathbb{V}$ نشان داده شود که $\mathbb{V}$ از کاردینالیتی K است. هدف ما به حداکثر رساندن آسیب مورد انتظار به اهداف است که معادل به حداقل رساندن ارزش هدف مورد انتظار است. باتوجه‌به $q_{ij} = (1 - P_{ij})$، خواهیم داشت:

$$\text{minimize } \sum_{j=1}^{N} V_j \prod_{i=1}^{M} q_{ij}^{x_{ij}} \qquad (۴)$$

به‌طوری‌که :

$$\sum_{j=1}^{N} x_{ij} = 1, \qquad i = 1,2,\dots,M$$

$$V_j \in \mathbb{V} \in \mathbb{R}_+^K, \qquad j = 1,2,\dots,N$$

$$x \in \mathbb{B}^{M\times N}$$

توجه داشته باشید که $\mathbb{R}$ نشان‌دهنده فضای اعداد واقعی، $\mathbb{V}$ و $\mathbb{B}$ به ترتیب به معنی عدد صحیح و زیر فضاهای دودویی است. محدودیت برابری[1] مسئلهٔ WTA نشان می‌دهد که هر جنگ‌افزار دقیقاً به یک هدف اختصاص داده شده است و هیچ‌چیز مانع این نیست که تمام جنگ‌افزارها به یک هدف واحد اختصاص داده شوند. فرمول‌بندی این مسئله برای اولین‌بار توسط [۴۲] پیشنهاد شد.

۳- روش پیشنهادی

جهت حل مسئلهٔ تخصیص جنگ‌افزار به هدف به‌صورت توزیع‌شده، برای هر یک از جنگ‌افزارها یک آتوماتای یادگیر در نظر می‌گیریم که می‌تواند به‌صورت غیرمتمرکز تصمیم‌گیری کند و باتوجه‌به توانایی خود جنگ‌افزار، شرایط محیط و شرایط هدف، یکی از اهداف را به خود تخصیص دهد. باتوجه‌به خصوصیات توزیع‌شدگی محیط، هیچ یک از جنگ‌افزارها از انتخاب دیگر جنگ‌افزارها اطلاع ندارد و بر اساس مقدار هدف باقی‌مانده و اثری که بر روی محیط دارد، از محیط پاداش دریافت می‌کند. هر آتوماتا به‌صورت جداگانه و باتوجه‌به بردار احتمال (p) مربوط به خودش عملی را از دامنهٔ مقادیر خود انتخاب می‌کند. هر LA، بعد از اینکه تمام LA ها انتخابشان را انجام دادند، پاداش دریافت می‌کند. پاداش باتوجه‌به اثری که در محیط گذاشته و بدون توجه به عمل دیگر عامل‌ها تعریف می‌شود؛ به این معنی که عامل باتوجه‌به میزان هدف باقی‌مانده در محیط پاداش دریافت می‌کند و این پاداش طبق الگوریتم یادگیری L_{RI}، که در بخش ۳ توضیح داده شد، می‌باشد. به‌عنوان مثال اگر تخصیص ایجاد شده در یک تکرار کوچکتر و یا مساوی طول تخصیص ایجاد شده تاکنون باشد و در یک تکرار، آتوماتای یادگیر j از مجموعه اقدام‌های مجاز خود اقدام i را انتخاب کرده باشد، احتمال انتخاب اقدام i طبق رابطه‌ی (۱) افزایش یافته و احتمال انتخاب سایر اقدام‌های آتوماتای یادگیر j باتوجه‌به رابطهٔ (۲) کاهش خواهد یافت.

در هر دور آتوماتاهای یادگیر (جنگ‌افزارها) باتوجه‌به پاداش‌های دریافتی، تابع یادگیری (p) خود را به‌روزرسانی می‌کنند و به سمت عمل بهینه پیش می‌روند. این تکرارها تا زمانی که دیگر، تمام جنگ‌افزارها نتوانند بهتر از قبل عمل کنند انجام می‌شود و نهایتاً به سمت همگرایی پیش می‌روند. p در دور آخر بهترین مسیر را مشخص می‌کند. بخش ۵ فرایند پیاده‌سازی مسئله را به‌خوبی نشان می‌دهد.

۳-۱- الگوریتم پیشنهادی

در این بخش نحوهٔ اجرای روش پیشنهادی را معرفی می‌کنیم. این الگوریتم‌ها نشان می‌دهند که چگونه مسئلهٔ تخصیص جنگ‌افزار به هدف توزیع‌شده با روش آتوماتای یادگیر حل می‌شود. الگوریتم ۱ الگوریتم اصلی می‌باشد؛ ابتدا متغیرهای اولیه مانند تعداد جنگ‌افزارها، تعداد اهداف، تعداد عمل هر جنگ‌افزار و همچنین پارامتر پاداش‌دهی، مقداردهی اولیه می‌شود (خطوط ۱. ۱ تا ۳. ۱). سپس به‌ازای تعداد عمل‌هایی که جنگ‌افزار می‌تواند انجام دهد، برای هر یک از هدف‌ها، یک مقدار احتمال ($1/r$) در نظر گرفته می‌شود (r تعداد عمل‌ها) (خط ۴. ۱).

در خط ۸. ۱ هر یک از عامل‌ها باتوجه‌به تابع تعریف شده در الگوریتم ۲ عمل خود را انتخاب می‌کنند، به این صورت که برای هر یک از عمل‌های موجود در مجموعه عمل‌ها، مجموع احتمالات محاسبه می‌شود (خط ۱. ۲ تا ۳. ۲)، اگر مقدار این مجموع کمتر از عدد تصادفی مدنظر باشد، همان عمل توسط عامل انتخاب می‌شود. (خطوط ۴. ۲ تا ۸. ۲).

Algorithm 1. Main

Input: Targets, $Weapon_{ij}$=the power of weapon *i* to destroying target *j*, *Targets*=value of targets, *itr*=number of iteration
Output: Assignment probabilities

1.1 initialize: *M*=number of Weapons, *N*=number of Targets,*r*=number of Actions, vector *totalTargetDestroy=0*
1.2 initialize *d=0.1* // reward parameter
1.3 initialize *b* = 0 // penalty parameter
1.4 for each action *k* a probability parameters p_{ik} is defined where is initially is set to $1/r$ *(in this paper we suppose r=N+1)*
1.5 repeat for each *itr*
1.6 initialize vector *A=0, effect =0*
1.7 for *i* in [1,2,..,*M*]
1.8 *selectAct* = **SelectAction(***r, M,* $Weapon_i$, *Actions,* p_i*)*
1.9 p_i = **Algorithm1 (***d, b, r, M,* $Weapon_i$, *Actions,* p_i, *selectAct,* $effect_i$*)*
1.10 endfor
1.11 for *i* in [1,2,..,*M*]
1.12 $effect_i$=$Weapon_{i,selectAct}$* $Target_{selectAct}$ // The amount of Target destroyed by the $Weapon_i$
1.13 end
1.14 end

[1] equality constraints

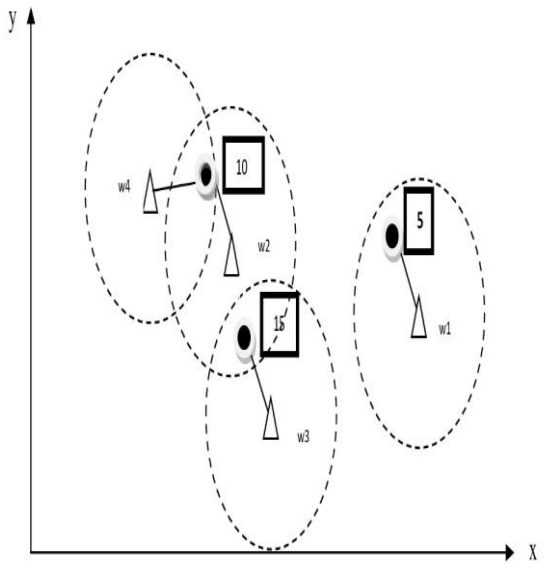
Algorithm 2. SelectAction

Input: *r, M,* $Weapon_i$*, Actions,* p_i

Output: action selection by Weapon: *selectAct*

2.1 initialize *sumP=0*
2.2 for *k* in [1,2,..,*r*]
2.3 $sumP \mathrel{+}= p_k$
2.4 if ($rand <= sumP$) // *rand* is a random number in [0,1]
2.5 $selectAct = k$
2.6 return *selectAct*
2.7 break
2.8 endif
2.9 endfor

در خط ۹. ۱ با استفاده از تابع تعریف شده در الگوریتم ۳ مقدار احتمالات p برای همهٔ عمل‌ها به‌روزرسانی می‌شود؛ این کار باتوجه‌به پاداشی که برای عامل‌ها بر اساس انتخابشان تعریف شده است، انجام می‌شود. یعنی اگر عامل‌ها در حالت جذب انتخاب-های بهتری از تکرار قبل داشته باشند و نتیجه تخصیص کار با استفاده از رابطه (۴) بهتر از تکرار قبل باشد، پاداش دریافت کرده و در غیر این صورت جریمه می‌شوند (خطوط ۱. ۳ تا ۶. ۳). الگوریتم ۳ برای به‌روزرسانی تابع احتمال از الگوریتم یادگیری L_{RI} مدل P که در رابطه (۱) و (۲) بیان شده است، استفاده می-کند(خطوط ۷. ۳ تا ۱۷. ۳).



شکل (۱). نمایشی از موقعیت جنگ‌افزارها و اهداف مهاجم در فضای x و y

Algorithm 3. UpdateProbability

Input: *d, b, r, M, Weapons, Actions, p, selectAct, effect, totalTargetDestroy*

Output: probability vector *p* after updating

3.1 *newtotalTargetDestroy* = sum(*effect*)
3.2 if *newtotalTargetDestroy* >= *totalTargetDestroy*
3.3 $\beta = 0$ //reward
3.4 else
3.5 $\beta = 1$ //penalty
3.6 endif
3.7 if $\beta == 0$
3.8 for *i* in [1,2,..,M]
3.9 for *k* in [1,2,..,r]
3.10 if *k* == *selectAct*
3.11 p_{ik} is updated by Eq.(1)
3.12 else
3.13 p_{ik} is updated by Eq.(2)
3.14 endif
3.15 endfor
3.16 endfor
3.17 endif
3.18 return *p*

۳-۲- شبیه‌سازی

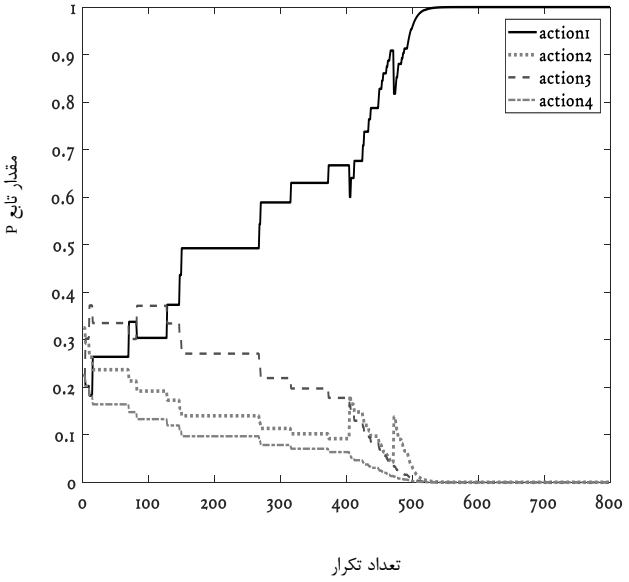
در تعریف ریاضی WTA، m جنگ‌افزار $W = \{W_1, W_2, \dots, W_m\}$ و n هدف مهاجم $T = \{T_1, T_2, \dots, T_n\}$ در یک محیط مانند شکل (۱) پراکنده شده‌اند. جنگ‌افزارها با شکل مثلث و هدف‌ها با دایره مشخص شده‌اند. محل کنونی جنگ‌افزار W_i با cur_pos_i و محل قرارگیری هدف T_j با cur_pos_j نشان داده می‌شود که در شکل (۱) در فضای x و y مشخص شده‌اند؛ فرض شده است که این موقعیت‌ها توسط جنگ‌افزارها شناخته شده است. همچنین مکان (یا موقعیت) جنگ‌افزارها و اهداف، در طول زمان در محیط ثابت است. هرجنگ‌افزار به‌صورت غیرمتمرکز عمل می‌کند و تنها می‌توانند اهداف را درون محدوده‌ی سنجش خود، حس کنند. عوامل ممکن است با انواع مختلف حسگرها، و در نتیجه محدوده سنجش ناهمگن مجهز شوند.

محدوده سنجش عامل A_i، توسط SR_i نشان داده می‌شود که در شکل (۱) با دایره‌های نقطه‌چین مشخص شده است. به خاطر محدودیت محدوده‌ی سنجش، هر عامل A_i می‌تواند تمام اهداف را درون یک فاصله SR_i از cur_pos_i مشاهده کند و نمی‌تواند هدفی را که دورتر است، مشاهده کند. همچنین فرض بر این است که هر جنگ‌افزار فقط می‌تواند یک هدف را به خود اختصاص دهد. هر کدام از جنگ‌افزارها می‌توانند به هر یک از اهداف تخصیص داده شوند. هر هدف دارای ارزش V_j است و بنا به این ارزش، تخصیص هدف با ارزش بیشتر، اهمیت بالاتری دارد. این ارزش‌ها در شکل (۱) با مربع‌های پررنگ نمایش داده شده است.

هر نوع جنگ‌افزار یک احتمال مشخص در جهت خرابی هر هدف دارد که با p_{ij} نشان داده می‌شود. در این فرمول، کاهش مقدار بقای مورد انتظار اهداف، همان جمع خرابی مورد انتظار است $q_{ij} = (1 - p_{ij})$. در این مسئله بیشتر از یک جنگ‌افزار می‌تواند به هر هدف اختصاص یابد و همهٔ اهداف نیازی به تخصیص همهٔ جنگ‌افزارها ندارند. تعداد جنگ‌افزارها از نوع i با x_{ij} مشخص می‌شود که x_{ij} تعداد جنگ‌افزارهای نوع i به هدف j ($\sum_{j=1}^{n} x_{ij} \le A_i \ for\ i = 1, \dots, n$) می‌باشد، اگر x_{ij} تعداد از جنگ‌افزارهای نوع i را به هدف j اختصاص داده شود، احتمال نجات هدف j به‌صورت $q_{ij}^{x_{ij}}$ خواهد بود. در مسئله WTA لازم است با اختصاص تعدادی از جنگ‌افزارهای نوع i به هدف j، جمع مقدار بقای مورد انتظار اهداف کاهش یابد. این مسئله می‌تواند توسط رابطه (۴) که در بخش ۳ گفته شد، نشان داده شود.

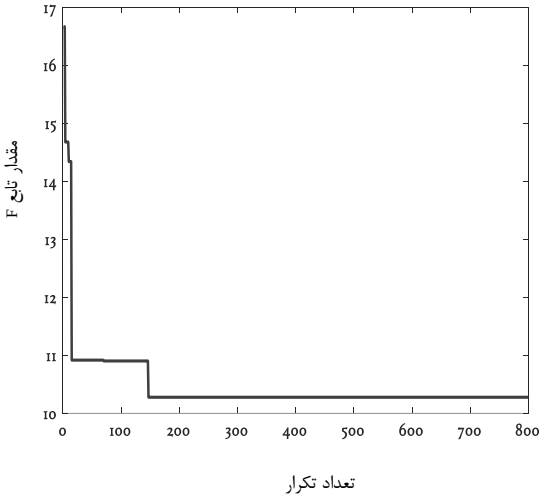
۴- نتایج

پیاده‌سازی ۱: در پیاده‌سازی اول با وجود ۸ جنگ‌افزار و ۳ هدف، برای هر یک از جنگ‌افزارها، ۴ عمل[1] وجود دارد: انتخاب «هدف ۱»، انتخاب «هدف ۲»، انتخاب «هدف ۳» و انتخاب «هیچ یک از هدف‌ها». شکل (۲) مقدار احتمال انجام هر یک از هدف‌ها توسط یکی از جنگ‌افزارها را در طول روند یادگیری نشان می‌دهد. همان‌طور که از شکل مشخص است بعد از اتمام اجرای برنامه، «عمل ۱» دارای احتمال بالاتری است و بنابراین در انتهای یادگیری، انتخاب جنگ‌افزار خواهد بود. این عمل همان عملی است که اگر جنگ‌افزار آن را نسبت به سایر هدف‌ها انجام دهد، در نتیجه‌ی جنگ‌افزارها سراسری اثر بیشتری خواهد داشت.



شکل (۲). مقدار احتمال انتخاب هدف‌ها توسط یکی از جنگ‌افزارها در روند یادگیری (۸ عامل و ۳ کار)

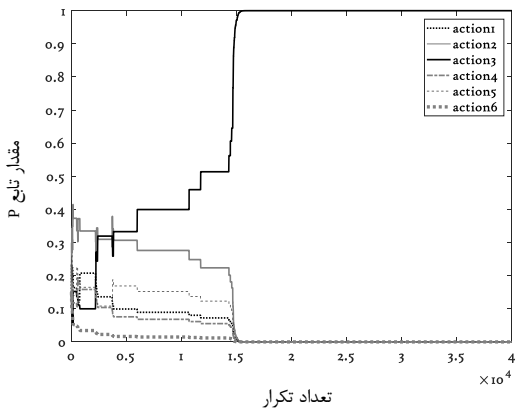
[1] action



شکل (۳). مقدار تابع F بعد از روند یادگیری (۸ عامل و ۳ کار)

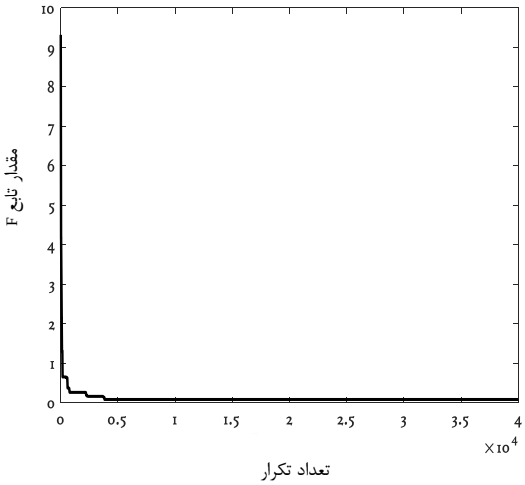
شکل (۳) نیز مقدار رابطهٔ (۴)، برای جنگ‌افزار را بعد از تکرارهای برنامه نشان می‌دهد که به سمت کمینه حرکت می‌کند. به‌خوبی مشخص است که پاداش‌ها به‌گونه‌ای تعریف شده است که جنگ‌افزارها به‌صورت توزیع‌شده و مجزا از هم بدون اطلاع از انتخاب یکدیگر و به جهت کمینه‌کردن تابع هدف، به سمت انتخاب درست حرکت می‌کنند.

پیاده‌سازی ۲: در فاز دوم پیاده‌سازی، مسئله برای تعداد جنگ‌افزار و هدف بیشتر انجام شده است. تعداد جنگ‌افزارها ۲۰ و تعداد هدف‌ها ۵ می‌باشد. این بار برای هر یک از جنگ‌افزارها ۶ عمل وجود خواهد داشت: انتخاب «هدف ۱»، انتخاب «هدف ۲»، انتخاب «هدف ۳»، انتخاب «هدف ۴»، انتخاب «هدف ۵» و انتخاب «هیچ‌یک از هدف‌ها». بنابراین احتمال به‌صورت $p = 1/6$ تعریف می‌شود. نتایج تغییرات p در شکل (۴) برای یکی از جنگ‌افزارها نشان داده شده است. باتوجه‌به این شکل مشخص است که بعد از اتمام روند یادگیری، جنگ‌افزار، عمل ۳ را انتخاب می‌کند، چرا که احتمال عمل ۳، به سمت ۱ میل می‌کند.



شکل (۴). مقدار احتمال انتخاب هدف‌ها توسط یکی از جنگ‌افزارها در روند یادگیری (۲۰ جنگ‌افزار و ۵ هدف)

شکل (۵)، مقدار تابع F بعد از اجرای الگوریتم، نشان می‌دهد که استفاده از آتوماتای یادگیر برای تعداد جنگ‌افزار زیاد هم توانسته نتیجه بهینه‌ای را برای تخصیص‌های مناسب هدف به دست بیاورد.



شکل (۵). مقدار تابع F بعد از روند یادگیری (۲۰ جنگ‌افزار و ۵ هدف)

پیاده‌سازی ۳: در فاز آخر پیاده‌سازی، باتوجه‌به همگرایی و زمان اجرای الگوریتم پیشنهادی در این مقاله برای موارد مختلف، نتایج را با یکی از روش‌های متمرکز –الگوریتم ژنتیک [۴۳، ۴۴]- و روش‌های پایه در مقاله‌ی [۴۵] که برای حل مسائل بهینه‌سازی محدودیت توزیع‌شده معرفی شده‌اند و از هیچ تکنیک یادگیری از نظر زمان اجرا و مقیاس‌بندی استفاده نمی‌کنند، مقایسه می‌کنیم. ستون اول، الگوریتم ژنتیک است که در این مقاله و باتوجه‌به رویکرد مسئله، پیاده‌سازی شده است. ستون دوم، الگوریتم-C DPOP، از روش معرفی شده در مقاله‌ی [۴۶] استفاده می‌کند. ستون سوم الگوریتم (LS-RAND) است که جستجوی محلی را با مقادیر اولیه تصادفی انجام می‌دهد. ستون آخر روش پیشنهادی است که در این مقاله به آن پرداخته شده است.

جدول (۱) زمان اجرای الگوریتم‌ها را باتوجه‌به تغییر در تعداد عامل‌ها و در نتیجه تعداد متغیرهای تصمیم‌گیری نشان می‌دهد. همان‌طور که در جدول مشاهده می‌شود، زمانی که تعداد عامل‌ها افزایش می‌یابد، زمان اجرای الگوریتم ژنتیک و الگوریتم -C DPOP به‌صورت نمایی افزایش می‌یابد. ستون سوم جدول نشان می‌دهد که LS-RAND بهتر از الگوریتم متمرکز و الگوریتم -C DPOP عمل می‌کند و در ستون آخر - الگوریتم پیشنهادی - به دلیل استفاده از یادگیری خودکار و حذف روش‌های مبتنی بر جستجو و روش اکتشافی، برای تعداد عامل زیاد مناسب است.

همان‌طور که از نتایج جدول (۱) مشخص است، رسیدن به تخصیص درست جنگ‌افزار به هدف، به‌خصوص زمانی که تعداد جنگ‌افزارها افزایش می‌یابد، با استفاده از روش متمرکز بسیار دیرتر از روش غیرمتمرکز می‌باشد و همین امر نشان می‌دهد که استفاده از روش غیرمتمرکز بسیار مناسب‌تر خواهد بود.

جدول (۱). زمان اجرای روش متمرکز و غیرمتمرکز جهت انجام تخصیص جنگ‌افزار به هدف

تعداد عامل‌ها	الگوریتم ژنتیک	الگوریتم C-DPOP	الگوریتم LS-RAND	الگوریتم پیشنهادی
	زمان (ms)	زمان (ms)	زمان (ms)	زمان (ms)
۲	۹۰۰	۲۲۳	۲۰۳	۰/۸
۴	۴۰۰۰	۴۸۹	۲۷۳	۱۰
۶	۱۰۳۴۱	۵۵۴۷	۳۸۵	۱۲/۱
۸	-	-	۵۵۶	۵۰/۳
۱۲	-	-	۱۰۹۲	۸۷
۱۶	-	-	۲۲۰۳	۱۲۱

۵- نتیجه‌گیری

در این مقاله یک روش مؤثر برای حل مسئلهٔ تخصیص جنگ‌افزار به هدف ارائه شده است که در آن از روشی غیرمتمرکز و توزیع‌شده استفاده شده است. این روش مبتنی بر آتوماتای یادگیر است و می‌تواند در یادگیری عامل‌ها برای تصمیم‌گیری صحیح و سریع، یاری کند. در واقع روش‌های متمرکز به‌درستی و بادقت و سرعت بیشتر نمی‌توانند این مسئله را حل کنند و این روش‌ها در زمان افزایش مقیاس مسئله، کارکرد عملیاتی خوبی ندارند؛ بنابراین استفاده از آتوماتای یادگیر به جهت اینکه روش نسبتاً ساده‌ای است و نیاز به اطلاعات کمی دارد، برای استفاده در محیط‌های چندعامله مناسب است.

پیاده‌سازی‌ها و نتایج نشان می‌دهند که در روشی که در این مقاله معرفی شده است، دیگر نیازی به استفاده از یک روش اکتشافی در کنار روش‌های مرسوم وجود ندارد و همین امر باعث کاهش زمان یادگیری خواهد بود. همچنین با افزایش مقیاس مسئله، روش معرفی شده در مقاله تابع هدف را در زمان کم، کمینه می‌نماید. بنابراین می‌توان از این روش در مواردی که نیاز به پاسخ بلادرنگ است، استفاده کرد.

۶- مراجع

[1] J. J. Enright, "*Efficient routing of multi-vehicle systems: limited sensing and nonholonomic motion constraint"s*. 2008.

[2] S. P. Lloyd and H. S. Witsenhausen, "Weapons allocation is NP-complete," in *1986 summer computer simulation conference*, 1986, pp. 1054-1058.

[3] P. A. Hosein, J. T. Walton, and M. Athans, "Dynamic weapon-target assignment problems with

1136-1146, 2007.
[20] M.-Z. Lee, "Constrained weapon–target assignment: Enhanced very large scale neighborhood search algorithm," *IEEE T SYST MAN CY A*, vol. 40, no. 1, pp. 198-204, 2009.
[21] J. M. Rosenberger, H. S. Hwang, R. P. Pallerla, A. Yucel, R. L. Wilson, and E. G. Brungardt, "The generalized weapon target assignment problem," Texas Univ at Arlington 2005.
[22] A. G. Kline, D. K. Ahner, and B. J. Lunday, "Real-time heuristic algorithms for the static weapon target assignment problem," *J HEURISTICS*, vol. 25, no. 3, pp. 377-397, 2019.
[23] Z. R. Bogdanowicz, "Advanced input generating algorithm for effect-based weapon–target pairing optimization," *IEEE T SYST MAN CY A*, vol. 42, no. 1, pp. 276-280, 2011.
[24] A. Tokgöz and S. Bulkan, "Weapon target assignment with combinatorial optimization techniques," *International journal of advanced research in artificial intelligence*, vol. 2, no. 7, pp. 39-50, 2013.
[25] F. Johansson and G. Falkman, "Real-time allocation of defensive resources to rockets, artillery, and mortars," in *2010 13th International Conference on Information Fusion*, 2010, pp. 1-8: IEEE.
[26] B. Xin, Y. Wang, and J. Chen, "An efficient marginal-return-based constructive heuristic to solve the sensor–weapon–target assignment problem," *IEEE T SYST MAN CYB: Systems*, vol. 49, no. 12, pp. 2536-2547, 2018.
[27] W. Jian and C. Chen, "Sensor-weapon joint management based on improved genetic algorithm," in *2015 34th Chinese Control Conference (CCC)*, 2015, pp. 2738-2742: IEEE.
[28] Z. T. Wang, H. J. Zhang, Y. Huang, K. Cheng, and T. Y. Wu, "Fire distribution optimization based on quantum immune genetic algorithm," in *2011 International Conference of Information Technology, Computer Engineering and Management Sciences*, vol. 1, pp. 95-98, 2011.
[29] F. Xue-yuan, X. Qing-hua, and T. Feng, "Dynamic target assignment base on ameliorated genetic algorithm," in *2010 International Conference on Computer Application and System Modeling (ICCASM 2010)*, vol. 15, pp. 443 -445, 2010.
[30] D. G. Galati, "Game theoretic target assignment strategies in competitive multi-team systems," University of Pittsburgh, 2005.
[31] S. Zeng, W. Wang, D. Ding, and Y. Zhang, "Target allocation method based on dynamic game," *Electronics Optics & Control*, vol. 18, no. 2, pp. 26-29, 2011.
[32] Y. Yang and X. Liu, "Task assignment based on improved dynamic contract net and ant colony search strategy," in *Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC)*, 2013, pp. 2880-2883: IEEE.
[33] R. Ghorbani Saber, M. Ranjbar, S. Balochian, and A. Izadipour, "Modelling and optimal solving of dependent sensor-weapon/threat assignment and scheduling problem by a metaheuristic algorithm based on GRASP," *Electronic and Cyber Defense*, vol. 8, no. 1, pp. 35-50, 2020 (in persian).
[34] M. Alighanbari and J. P. How, "Decentralized task vulnerable C2 nodes", Massachusetts Inst of Tech Cambridge Lab for Information and Decision Systems, 1988.
[4] P. A. Hosein and M. Athans, "Preferential defense strategies", pp.1-25, 1990.
[5] P. A. Hosein and M. Athans, "Some analytical results for the dynamic weapon-target allocation problem", MASSACHUSETTS INST OF TECH CAMBRIDGE LAB FOR INFORMATION AND DECISION SYSTEMS, 1990.
[6] S. Han, "Stochastic theory and method for effectiveness analysis of missile weapon systems", *BeiJing National Defense Industry Press*, pp. 77-78, 2001.
[7] Z. J. Lee, C. Y. Lee, and S. F. Su, "A Hybrid Genetic Algorithm Applied to Weapon-Target Assignment Problem", 2004.
[8] D. Khosla, "Hybrid genetic approach for the dynamic weapon-target allocation problem", in *Battlespace digitization and network-centric warfare*, pp. 4396, pp. 244-259: International Society for Optics and Photonics.
[9] R. K. Ahuja, J. B. Orlin, and A. Tiwari, "A greedy genetic algorithm for the quadratic assignment problem," *COMPUT OPER RES*, vol. 27, no. 10, pp. 917-934, 2000.
[10] Z. R. Bogdanowicz, A. Tolano, K. Patel, and N. P. Coleman, "Optimization of weapon–target pairings based on kill probabilities," *IEEE transactions on cybernetics*, vol. 43, no. 6, pp. 1835-1844, 2012.
[11] K. Volle, J. Rogers, and K. Brink, "Decentralized cooperative control methods for the modified weapon–target assignment problem," *J GUID CONTROL DYNAM*, vol. 39, no. 9, pp. 1934-1948, 2016.
[12] H. Cai, J. Liu, Y. Chen, and H. Wang, "Survey of the research on dynamic weapon-target assignment problem," *J SYST ENG*, vol. 17, no. 3, pp. 559-565, 2006.
[13] J. Chen, J. Yang, and G. Ye, "Auction algorithm approaches for dynamic weapon target assignment problem," in *2015 4th International Conference on Computer Science and Network Technology (ICCSNT)*, vol. 1, pp. 402-405, 2015.
[14] H. Alimohammady and V. Tabatabavakily, "A Novel Algorithm for Radar Jamming Resource Allocation", *Electronic and Cyber Defense*, vol. 7, no. 3, pp. 53-70, 2019.
[15] K. Tuyls, "Learning in Multi-agent Systems. An Evolutionary Game Theoretic Approach," 2004.
[16] M. A. Thathachar and P. S. Sastry, *Networks of learning automata: Techniques for online stochastic optimization*. Springer Science & Business Media, 2003.
[17] M. D. Pendrith and M. J. McGarity, "An analysis of non-Markov automata games: Implications for reinforcement learning." *University of New South Wales, School of Computer Science and Engineering*, 1997.
[18] P. Vrancx, K. Verbeeck, and A. Nowé, "Networks of learning automata and limiting games," in *Adaptive Agents and Multi-Agent Systems III. Adaptation and Multi-Agent Learning*: Springer, 2005, pp. 224-238.
[19] R. K. Ahuja, A. Kumar, K. C. Jha, and J. B. Orlin, "Exact and heuristic algorithms for the weapon-target assignment problem," *OPER RES*, vol. 55, no. 6, pp.

[40] A. R. Eckler and S. A. Burr, "Mathematical Models Of Target Coverage And Missile Allocation," 1972.
[41] S. Matlin, "A Review of the Literature on the Missile-Allocation Problem," *OPER RES*, vol. 18, no. 2, pp. 334-373, 1970.
[42] A. S. Manne, "A Target-Assignment Problem," *OPER RES*, vol. 6, no. 3, pp. 346-351, 1958.
[43] M. A. Sahin and K. Leblebicioglu, "A genetic algorithm for weapon target assignment problem," *presented at the Proceedings of the 2009 Summer Computer Simulation Conference, Istanbul, Turkey*, 2009.
[44] S. Mirjalili, "Genetic Algorithm," in *Evolutionary Algorithms and Neural Networks: Theory and Applications* Cham: Springer International Publishing, 2019, pp. 43-55.
[45] K. D. Hoang, F. Fioretto, P. Hou, M. Yokoo, W. Yeoh and R. Zivan. "Proactive dynamic distributed constraint optimization". *In: Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2016, pp. 597–605.
[46] A. Petcu and B. Faltings. "DPOP: A scalable method for multiagent constraint optimization". *In IJCAI 05*, pp. 266-271, 2005.
assignment for unmanned aerial vehicles," in *Proceedings of the 44th IEEE Conference on Decision and Control*, 2005, pp. 5668-5673: IEEE.
[35] S.-Y. Tang, S. Mei, Y.-F. Zhu, Y.-L. Lei, and Q. Li, "Distributed weapon target assignment algorithm based on extended contract net protocol," *Systems Engineering and Electronics*, vol. 33, no. 3, pp. 568-574, 2011.
[36] L. Cao and A. Zhang, "Cooperative target assignment for unmanned combat aerial vehicles based on Bayesian optimization algorithm with decision graphs," in *2013 Ninth International Conference on Natural Computation (ICNC)*, pp. 439-443, 2013.
[37] Y. Chen, H. Cai, and L. Xing, "An improved algorithm of policies optimization of dynamic weapon target assignment problem," *Systems Engineering-Theory & Practice*, vol. 7, pp. 160-165, 2007.
[38] L. Wang, H. Zheng, and X. Zheng, "Survey on resource-constrained project scheduling under uncertainty [J]," *Control and Decision*, vol. 29, no. 4, pp. 577-584, 2014.
[39] K. S. Narendra and M. A. L. Thathachar, *Learning Automata: An Introduction*. Dover Publications, 2013.