Military Science and Tactics
AJA Command and Staff University
https://www.qjmst.ir/
Online ISSN: 4520-2676
Print ISSN: 2008-191x

Proposing a Conceptual Model for the Application of Artificial Intelligence Algorithms in Offensive Cyber Operations

Ahmad Reza Tarasoli[1]| Hasan Mohammadi Monfared[2]✉

1. Doctoral student of Cyber Department, Faculty of National Security, Higher National Defense University, Tehran, Iran.
Email: a. tarrasoli01@sndu. ac. ir
2. Corresponding author, Assistant Professor of Cyber Department, Faculty of National Security, Higher National Defense University, Tehran, Iran.
E-mail: h. mohammadi@sndu. ac. ir

Article Info

Article type:
Research Article

Article history:
Received
4 March 2024
Received in revised form
23 October 2024
Accepted
10 December 2024
Published online
14 March 2025

Keywords:
Artificial Intelligence, Artificial Intelligence Algorithm, Offensive Cyber Operations, Miter Attack Framework, Cyber Kill Chain

ABSTRACT

Objective: The primary objective of this study is to develop a conceptual model for the application of artificial intelligence (AI) algorithms in cyber operations, grounded in the MITRE ATT&CK framework. The model aims to enhance understanding of how AI can be systematically exploited to support offensive cyber activities.

Methodology: In order to achieve the mentioned goal, first and in the stage of data acquisition by initial search in Google, Google Scholar and reliable scientific databases (Science Direct, Springer, Emerald, Taylor, and Francis) articles related to artificial intelligence and offensive cyber operations are explored and in general, 121 results were found. Then, using the text theme analysis approach, 39 codes were extracted in the form of basic themes.

Findings: The thematic network revealed that AI technologies can significantly influence various stages of cyber operations. Specifically, the organizing themes—identification, implementation, anonymity, and impact and action—highlight the diverse functionalities AI provides across the cyber-attack lifecycle. These include target recognition, adaptive strategy formulation, concealment of operational signatures, and maximization of offensive outcomes.

The finalized conceptual model articulates the operational domains where AI algorithms are most effectively utilized within the context of the MITRE ATT&CK framework. It demonstrates the potential of AI to enhance adversarial tactics, techniques, and procedures (TTPs), particularly by enabling automated decision-making, increasing situational responsiveness, and improving operational precision.

Conclusion: The findings underscore that the integration of artificial intelligence into cyber operations substantially reduces dependence on human operators while automating complex stages of the cyber-attack process. AI algorithms enhance the accuracy, speed, and overall effectiveness of cyberattacks, thereby transforming the nature of offensive cyber capabilities.

Cite this article: Tarasoli, Aamadreza, & Mohammadimonfared, Hasan (2024). Presenting a conceptual model of the application of artificial intelligence technology algorithms in Offensive cyber operations. *Military Sciences & Techniques*, 20(70), 209- 240.
DOI: http//doi. org/ 10. 22034/qjmst. 2025. 2024258. 2031

Publisher: AJA Command and Staff University
DOI: 10. 22034/qjmst. 2025. 2024258. 2031

Military Science and Tactics
https://www.qjmst.ir/ Online ISSN: 4520-2676 Print ISSN: 2008-191x

ارائه الگوی مفهومی کاربست الگوریتم‌های فناوری هوش مصنوعی در عملیات سایبری تهاجمی

احمدرضا ترسلی[1] | حسن محمدی‌منفرد[2]✉

۱. دانشجو دکتری گروه سایبر، دانشکده امنیت ملی، دانشگاه عالی دفاع ملی، تهران، ایران، رایانامه: a. tarrasoli01@sndu. ac. ir

۲. نویسنده مسئول، استادیار گروه سایبر، دانشکده امنیت ملی، دانشگاه عالی دفاع ملی، تهران، ایران، رایانامه: h. mohammadi@sndu. ac. ir

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی

تاریخ دریافت:
۱۴۰۲/۱۲/۱۴

تاریخ بازنگری:
۱۴۰۳/۰۸/۰۲

تاریخ پذیرش:
۱۴۰۳/۰۹/۲۰

تاریخ انتشار:
۱۴۰۳/۱۲/۲۴

کلیدواژه‌ها:
هوش مصنوعی، الگوریتم هوش مصنوعی، عملیات سایبری تهاجمی، چارچوب حملات مایتره، زنجیره کشتن سایبری

چکیده

هدف: ارائـه الگوی مفهومی به‌منظور بهره‌برداری از الگوریتم‌های هوش مصنوعی در عملیات سایبری، مبتنی بر چارچوب حملات مایتره اســت.

روش‌شناسی: بــرای دسـتیابی بـه هـدف مذکور، نخست و در مرحله داده‌یابی با جستجوی اولیه در گوگل، گوگل‌اسکولار و پایگاه‌های علمی معتبر (ساینس دایرکت، اسپرینگر، امرالد، تیلور و فرانسیس) نوشته‌هایی در پیونـد بـا موضـوع هوش مصنوعی و عملیات سایبری تهاجمی کاوش و در مجموع، ۱۲۱ نتیجه یافت شد. سپس با استفاده از رویکرد تحلیل مضمون متنی، تعداد ۳۹ کد در قالب مضامین پایه استخراج شد.

یافته‌ها: در ادامه و پس از تلفیق و ترکیب کدهای مختلف، شبکه مضامین پژوهش در قالب ۱۸مضمون پایه و چهار مضمون سازمان دهنده مشتمل بر (شناسایی، اجرا، گمنامی، ضربه و اقدام) استخراج گردید. با مراجعه به جامعه خبرگی تحقیق، در قالب یک پنــل دلفــی متشــکل از ۱۰ نفر متخصــص، آزمون کیفیت و اعتبارسنجی دسته‌بندی‌ها صورت گرفت. ســرانجام و پس از اعمال نظرات جامعه خبرگی پژوهش، الگوی مفهومی کاربسـت الگوریتم‌های فناوری هوش مصنوعی در عملیات سایبری ترسیم شــد.

نتیجه گیری: تبیین نتایج تحقیق نشان می‌دهد که الگوریتم‌های هوش مصنوعی، ضمن کاهش نقش عامل انسانی و خودکارسازی مراحل حمله، دقت، سرعت و اثربخشی حملات را افزایش چشمگیری می‌دهند.

استناد: ترسلی، احمدرضا؛ و محمدی منفرد، حسن (۱۴۰۳). ارائه الگوی مفهومی کاربست الگوریتم‌های فناوری هوش مصنوعی در عملیات سایبری تهاجمی. *علوم و فنون نظامی*، ۲۰(۷۰)، ۲۴۰- ۲۰۹.

DOI: http//doi. org/ 10. 22034/qjmst. 2025. 2024258. 2031

ناشر: دانشگاه فرماندهی و ستاد ارتش جمهوری اسلامی ایران

DOI: 10. 22034/qjmst. 2025. 2024258. 2031

Military Science and Tactics
https://www.qjmst.ir/

Proposing a Conceptual Model for the Application of Artificial Intelligence Algorithms in Offensive Cyber Operations

Aamadreza Tarasoli[1] | **Hasan Mohammadimonfared**[2]✉

Extended Abstract

Introduction

With the increasing development of artificial intelligence technology and the increase in the number, intensity and variety of offensive cyber operations in recent years (Walker-Munro, et al. ,2023: 28), cyber attackers have turned to using advanced cyberattack techniques to Automate cyber weapons with the help of artificial intelligence and machine learning (LeCun, 2022: 1), increase their accuracy and generally reduce the dependence on human power (Nguyen, et al. , 2023: 1). Offensive cyber operations can be defined as "actions that affect the confidentiality, integrity, and availability of information" (Huskaj,2019: 1). A well-known example of a potential govern ment offensive cyber operation is the 2020 supply chain attack that compromised tens of thousands of government computers. The attack was one of the worst government breaches in US history (FireEye,2023). Considering the role of artificial intell igence technology and machine learning algorithms to automate the exploitation of cyberspace and the successful implementation of cyberattacks (Maeda, et al. ,2021: 11). Therefore, the main goal and concern of the researcher is to present a conceptual model in order to use artificial intelligence technology algorithms in

[1] Doctoral student of Cyber Department, Faculty of National Security, Higher National Defense University, Tehran, Iran.
Email: a.tarrasoli01@sndu. ac. Ir
[2] Corresponding author, Assistant Professor of Cyber Department,Faculty of National Security, Higher National Defense University, Tehran, Iran.
Email: h. mohammadi@sndu. ac. Ir

offensive cyber operations, based on the framework of Maître attacks.

Methodology

In terms of direction, this research is considered to be of the type of applied development studies, and in terms of the nature of the data, it is of the qualitative research type, and in terms of the certainty of the data, it is of the exploratory type. The method of information analysis is text content analysis. Thematic analysis is a method to recognize, analyze and report patterns (themes) in qualitative data. This method is a process for analyzing textual data and transforms scattered and diverse data into rich and detailed data (Abdi Jafari et al. , 2012: 153). Themes are extracted from within the codes and developed by interpretive analysis of the data; Overarching themes form the core and focus of a the- matic network. (ibid: 154). The content validity test was used to measure the validity of the research. Validity refers to the fact that; To what extent do the obtained data adequately reflect the true meaning of the concept under consideration (Babi, 2012: 335). In fact, validation means “do we really measure what we intend to measure”(Baker, 2014: 138). Therefore,in the leading research,in order to validate the findings,the judgment and opini- on of ten researchers of the research field, consisting of faculty members and researchers of the National Defense Uni versity, have been used in the form of a Delphi panel. Also, to measure the reliability in this research, the agreement percentage index (Cohen's kappa coefficient) has been used.

In this way, after the end of coding, about 20% of codings were selected and coded by the researchers again. The frequency of code similarity was calculated using this index, which was determined to be higher than 70% for all components.

Findings

In this research, a total of 121 results were found in the initial search in Google, Google Scholar and authoritative scientific database(Science Direct, Springer, Emerald, Taylor, and Francis) by searching the words artificial intelligence technology and offensive cyber operations. In the first stage, 48 sources were excluded because they were not directly related to the topic; In the second stage, 17 duplicate sources were removed from the list of sources and in the final stage, 56 sources were carefully studied to extract the network of themes, and themes and codes were extracted from 27 sources and from 29 other sources due to overlapping with other sources. No topics were found to extract. Based on this,finally, 27 sources were selected for in-depth study as described in Table (1).

Table (1) number of selected articles from scientific databases

scientific base	Science Direct	Springer	Emerald	Taylor and Francis	Arxio	Magiran	total
number	10	4	6	2	3	2	27

After finishing the initial coding and counting 39 codes, common codes were put together to decide how to combine and combine different codes to form basic themes. In the next step, the nature and concept of each basic theme was identified more precisely and appropriate names were chosen for the themes. Then, higher level themes (organizing themes) were identified. For this purpose, first, the main themes that had a lot in common or were

Military Science and Tactics
https://www.qjmst.ir/

around a specific topic were combined and formed an organizing theme. The results of this analysis include the identification of the basic theme, the organizing theme, and the overarching theme in the form of a conceptual model of the use of artificial intelligence algorithms in offensive cyber operations according to Figure (1).

Figure(۱): conceptual model of using artificial algorithms in offensive cyber operations

Conclusion

Basically, cyber attackers, whether governmental or non-governmental actors, are constantly changing and improving the efficiency of their attacks and emphasize the use of artificial intelligence algorithms in the attack process. This study identifies and introduces the offensive capabilities of artificial intelligence that allow attackers to launch attacks on a larger scale, with a wider scope, faster and more precisely.

This research showed that in addition to the widespread use of artificial intelligence algorithms in all 14 steps of the Maître attack framework, most algorithms are used in the stages of identification, access promotion, discovery, lateral movement and access to credentials. Also, according to the results of the interviews, despite the fact that the participants believed that there are certain complications in knowing the algorithms of artificial intelligence in cyber operations and sharing information in this field, it was confirmed that the use of artificial intelligence algorithms It can bring various benefits such as the following:

1. Increasing the accuracy and efficiency of cyber operations
2. Reducing the cost and time of cyber operations
3. Achieving more complex goals

However, the use of artificial intelligence algorithms in offensive cyber operations also brings challenges that need to be used with awareness and caution. It should be possible to be sure that artificial intelligence algorithms will perform well in different conditions and provide predictable results. They must be transparent and accurately identify targets, execute attacks effectively, and responsibly avoid computational errors.

References

- Abedi Jafari, Hassan; and others (2013), theme analysis and theme network: a simple and efficient method to explain patterns in qualitative data, *Strategic Management Thought Quarterly*, fifth year, number 2.
- Baker, Therese L (2012). *How to conduct social research*. Translated by Houshang Naibi. Tehran: Ney Publishing.
- Bebi, Earl (2012), *research methods in social sciences: the first volume*. Translated by Reza Fazel. Tehran: Organization for the study and compilation of university humanities books (Samt), 9th edition.
- FireEye, 2023, *Highly Evasive Attacker Leverages SolarWinds Supply Chain to Compromise Multiple Global Victims With SUNBURST Backdoor*. Available at: https: //www. fireeye.

Online ISSN: 4520-2676 Print ISSN: 2008-191x

com/blog/threat-research/2020/12/evasive-attacker-leveragessolarwinds-supply-chain-compromises-with-sunburst-backdoor. html

- Huskaj, G. , (2019), *The current state of research in offensive cyberspace operations. In 18th European Conference on Cyber Warfare and Security* (ECCWS 2019), 4-5 July 2019, Coimbra, Portugal (pp. 660-667). Academic Conferences and Publishing International Limited.
- LeCun, Y. , (2022), A path towards autonomous machine intelligence, *Courant Institute of Mathematical Sciences*, New York University.
- Maeda, R. , & Mimura, M. , (2021), *Automating post-exploitation with deep reinforcement learning. Computers & Security*.
- Mirsky, Y. , Demontis, A. , Kotak, J. , Shankar, R. , Gelei, D. , Yang, L. ,... & Biggio, B. , (2023), *The threat of offensive ai to organizations, Computers & Security*.
- Nguyen, H. V. , & Uehara, T. , (2023), Multilayer Action Representation based on MITRE ATT&CK for Automated Penetration Testing. *Journal of Information Processing*.
- Walker-Munro, B. , Mount, D. , & Ioannou, R. , (2023), *The Hacker Strikes Back: Examining the Lawfulness of "Offensive Cyber" Under the Laws of Australia*.

مقدمه

با پیشرفت روزافزون فناوری هوش مصنوعی و افزایش تعداد، شدت و تنوع عملیات‌های سایبری تهاجمی[1] در سال‌های اخیر (Walker-Munro, et. al, 2023: 28)، مهاجمان سایبری به استفاده از فن‌های پیشرفته حمله سایبری روی آورده‌اند تا سلاح‌های سایبری[2] را به کمک هوش مصنوعی و یادگیری ماشین خودکار کرده (LeCun, 2022: 1)، دقت آن‌ها را افزایش داده و در کل وابستگی به نیروی انسانی را کاهش دهد (Nguyen, et. al, 2023: 1). عملیات سایبری آفندی را می‌توان به‌عنوان "اقداماتی که بر محرمانگی، یکپارچگی و دردسترس‌بودن اطلاعات تأثیر می‌گذارد" تعریف کرد (Huskaj, 2019: 1). به طور خاص، عملیات سایبری به‌عنوان حملات سازمان‌یافته علیه دارایی‌های دیجیتالی نهادهای خاص انجام می‌شوند. بازیگران مختلفی ممکن است عملیات سایبری را انجام دهند. نمونه‌ای از این عملیات، در فضای سایبری تحت رهبری دولت‌ها و باهدف دسترسی به سامانه‌های حیاتی دشمن انجام می‌شود. یک نمونه شناخته شده از عملیات سایبری آفندی احتمالاً دولتی، حمله زنجیره تأمین[3] در سال ۲۰۲۰ است که منجر به آسیب‌دیدن ده‌ها هزار رایانه دولتی شد. این حمله یکی از بدترین تخلفات دولتی در تاریخ ایالات متحده بود (FireEye, 2023). شناسایی دقیق آسیب‌پذیری‌ها در فضای سایبری و سپس بهره‌برداری از آنها یکی از مهم‌ترین مسائل پیشروی در عملیات سایبری است که به دلایل عدیده‌ای چون تنوع زیرساخت‌های نرم‌افزاری و سخت‌افزاری، به‌روزرسانی و دریافت مستمر وصله‌های امنیتی، کشف آسیب‌پذیری‌های موجود را برای عامل انسانی بسیار سخت و در مواردی غیرممکن ساخته است.

از دیگر مسائل مهم پیشروی عملیات سایبری تهاجمی، شناخت تاکتیک‌های فریب دشمن چون کوزه عسل[4] و شبکه عسل[5]، توسط نیروی انسانی مجری و هدایت‌کننده عملیات است. از طرف دیگر رعایت اصول پیچیدگی، گمنامی و تشخیص زمان مناسب از چالش‌های پیشرو برای اجرای عملیات موفق است. چراکه؛ اجرای دقیق عملیات سایبری

[1] Offensive Cyber Operations
[2] Cyber warfare
[3] SolarWinds
[4] Honeypot
[5] Honeynet

به دلیل محدودیت‌های شناختی عامل انسانی، اغلب منجر به تحمیل زمان و هزینه قابل‌توجه گردیده و در مواردی زمینه‌ساز شکست‌های حفاظتی را فراهم می‌نماید. باتوجه‌به نقش هوش مصنوعی و الگوریتم‌های یادگیری ماشین برای خودکارسازی بهره‌برداری از فضای سایبری و اجرای موفق حملات سایبری (Maeda,et. al, 2021: 11). مسائل و دغدغه‌های فراوانی را برای اطمینان از چگونگی به‌کارگیری مناسب هوش مصنوعی در عملیات سایبری ایجاد خواهد کرد. دغدغه‌هایی که حاکی از فقدان چارچوب‌های لازم برای اطمینان از توسعه و استفاده مفید از هوش مصنوعی است که خود از جمله مسائل مهم و اولویت‌دار محسوب می‌گردد.

در خصوص ضرورت بحث نیز می‌توان به چند نکتهٔ مهم اشاره داشت که استفاده از الگوریتم‌های هوش مصنوعی در عملیات سایبری آفندی چالش‌هایی نیز به همراه دارد که نیاز است در استفاده از آن‌ها با آگاهی و احتیاط عمل کرد. باید بتوان اطمینان داشت که الگوریتم‌های هوش مصنوعی در شرایط مختلف عملکرد مناسبی خواهند داشت و نتایج قابل‌پیش‌بینی ارائه خواهند کرد. آن‌ها بایستی شفاف بوده و اهداف را به طور دقیق شناسایی کنند، حملات را به طور مؤثر اجرا کنند و به طور مسئولانه از خطاهای محاسباتی جلوگیری کنند. براین‌اساس سؤال اصلی و دغدغه محقق این است که کاربرد هوش مصنوعی در عملیات سایبری تهاجمی چیست؟ و چگونه می‌تواند بر میزان دانش و درک مدیران حوزه امنیتی و دفاعی کشور در حوزه ارتقای توان رزم سایبری بیفزاید؛ لذا؛ در این تحقیق تلاش گردید تا با ارائه الگوی مفهومی کاربست الگوریتم‌های فناوری هوش مصنوعی در عملیات سایبری، ضمن ارتقای توان رزم سایبری در سازمان‌های مرتبط در کشور، زمینه علمی و دانش لازم جهت بازدارندگی سایبری نیز فراهم گردد. البته در این زمینه ابتکارها و الزامات اساسی دیگری نیز چون جذب و نگهداشت نیروی انسانی متخصص، تأمین اعتبارات و زیرساخت‌های پردازشی لازم جهت توسعه کاربرد هوش مصنوعی در آینده ضروری و اجتناب‌ناپذیر خواهد بود که در این مقاله به آن ورود نشده است.

مبانی نظری و پیشینه‌شناسی تحقیق

پیشینه تحقیق

(۱) قاسمی و همکاران (۱۴۰۲)، در تحقیقی با عنوان «الگوی ارزیابی قدرت آفند سایبری ارتش جمهوری اسلامی ایران» نشان دادند که نیروی انسانی، پیچیدگی سایبری، تسلیحات

سایبری و آگاهی وضعیتی سایبری مهم‌ترین مولفه‌ها در آفند سایبری هستند و در نهایت ۱۲ شاخص برای ارزیابی قدرت آفند سایبری ارتش جمهوری اسلامی ایران را ارائه نموده‌اند.

(۲) رمضان‌زاده و همکاران (۱۳۹۹) در تحقیقی با عنوان «ارائه مدل مفهومی ارزیابی قدرت سایبری نیروهای مسلح با تاکید بر بعد بازدارندگی سایبری» نشان دادند که مدل مفهومی دارای مولفه‌های پنج‌گانه پشیمان کنندگی دشمن، استمرار عملیات، پاسخ به تهاجم، استحکام‌سازی و بازیابی است. همچنین، مولفه استمرار عملیات، از اولویت بالاتری نسبت به سایر مولفه‌ها برخوردار است.

(۳) وایت، کریستوفر[1] (۲۰۲۰)، در تحقیقی با عنوان «مشکلات سمی: پارادایم‌های جدید و رقابت "توافق شده" در عصر عملیات سایبری مجهز به هوش مصنوعی» نشان داد که: فناوری‌های هوش مصنوعی عملیات سایبری آفندی را تقویت می‌کنند. او با اشاره به سیاست تعامل مداوم، رقابت مورد توافق و دفاع رو به جلو که در سال ۲۰۱۸ توسط ایالات متحده اعلام شد، بحثی در مورد پیامدهای بازدارندگی در فضای سایبری ارائه می‌کند.

(۴) سرهنگ شانکار،[2] آرون، (۲۰۲۳)، در تحقیقی با عنوان «عملیات تهاجمی فضای مجازی: استفاده از هوش مصنوعی و زنجیره‌های کشتن[3] برای تجزیه و تحلیل اثرات مرجع اجرایی» مشخص نمود که: اختیارات عملیات سایبری تهاجمی را می‌توان به طور مسئولانه و مؤثر به فرماندهان تفنگداران دریایی[4] تفویض کرد. جنگ آینده مستلزم قابلیت های منظم جنگ سایبری است و فرماندهان راهکنشی نیاز به اختیار عمل دارند. مدل‌های هوش مصنوعی برای بهینه‌سازی این چالش تصمیم گیری وجود دارد. از یک الگوریتم هوش مصنوعی نظارت شده می توان برای ارتقای تفویض اختیار عملیات سایبری به فرماندهی تفنگداران دریایی استفاده کرد و تغییرات لازم را در قانون و خط‌مشی برای دستیابی به آن هدف بیشتر برجسته کرد.

(۵) هوسکاج، گازمند[5]، (۲۰۱۹)، در تحقیقی با عنوان «وضعیت کنونی تحقیق در عملیات تهاجمی فضای مجازی» نشان داد که: بعد از استفاده از سلاح سایبری استاکس

[1] Christopher White
[2] LtCol Shankar, Arun
[3] kill chains
[4] Marine Air-Ground Task Force (MAGTF)
[5] Huskaj, Gazmend

نت در سال ۲۰۱۰ علیه سانتریفیوژهای غنی‌سازی هسته‌ای ایران، دولت‌های ملی در حال توسعه قابلیت های فضای سایبری برای انجام عملیات های آفندی فضای سایبری هستند. و در نتیجه، تحقیقات در مورد عملیات سایبری آفندی با عمومی شدن اطلاعات بیشتر در حال بلوغ است. ثانیاً، تحقیقات کنونی برخی از ایده‌های اساسی خوب را در مورد اثراتی که می توان از طریق عملیات سایبری آفندی به دست آورد، نحوه انجام آن‌ها و ابزارها، فن‌ها و رویه‌های مرتبط را فهرست می‌کند.

پذیرفتن بهره‌گیری از دانش نهفته در فناوری هوش‌مصنوعی در محیط پیچیده امروزی برای حفظ برتری و آمادگی سازمان‌ها در فضای سایبر به‌عنوان قلمرویی قدرت ساز، از جمله مواردی می‌باشد که در تحقیقات انجام‌شده قبلی موردتوجه قرارگرفته است. همچنین، تبین ابعاد و یکپارچگی الگوی مفهومی کاربست الگوریتم‌های هوش مصنوعی از حیث کاربردپذیری در حوزه عملیات سایبری، از وجوه افتراق این تحقیق با تحقیقات انجام شده قبلی است. لذا؛ با توجه به اینکه تاکنون تحقیقی با این عنوان انجام نشده است از این رو پژوهش حاضر با شناسائی انواع الگوریتم‌های موثر در این زمینه، دارای نوآوری می‌باشد.

مبانی نظری تحقیق

عملیات سایبری آفندی: مأموریت‌هایی هستند که برای ایجاد قدرت در فضای سایبری از طریق اقداماتی در حمایت از اهداف ملی انجام می‌گردند. این عملیات، ممکن است به طور انحصاری عملکردهای فضای سایبری دشمن را هدف قرار دهد یا در حوزه‌های فیزیکی بر سامانه‌های تسلیحاتی، فرماندهی و واپایش، پشتیبانی و سایر اهداف باارزش تأثیر بگذارد. (JP 3-12, 2018: 6). قابلیت‌های تهاجمی در فضای سایبری از ابزارهای ساده فناوری که می‌توانند به‌سرعت توسعه یابند تا ابزارهای پیچیده‌ای که به یک دوره توسعه طولانی نیاز دارند، متغیر است. سطح فناوری موردنیاز عمدتاً به اثرات هدف، سخت‌شدن و پیچیدگی محیط هدف بستگی دارد. علاوه بر این، زمان آماده‌سازی در عملیات سایبری، ممکن است به بلوغ قربانی[1]، تلاش‌های جمع‌آوری اطلاعات، الزامات ناشناس‌سازی[2] و هرگونه اقدامات کاهشی موردنیاز برای واپایش خسارت جانبی بستگی داشته باشد (Pecoraro, et. al, 2021: 2).

[1] victim maturity
[2] anonymization requirements

مدل‌های رایج عملیات سایبری تهاجمی

سه مدل رایج که بیشترین ارجاع و استفاده را در عملیات سایبری تهاجمی داشته‌اند به ترتیب، زنجیره کشتن سایبری لاکهید مارتین[۱]، چارچوب حملات مایتره[۲] و مدل الماس[۳] به شرح زیر معرفی می‌گردند:

زنجیره کشتن سایبری لاکهید مارتین

عملیات پیچیده و ساختاریافتهٔ سایبری را می‌توان همانند زنجیره کشتن سایبری[۴]، به‌عنوان مجموعه‌ای از رویدادهای منظم برای هر عملیات سایبری چارچوب‌بندی کرد (Pecoraro, et. al, 2021: 2). زنجیره کشتن سایبری یک رویکرد سامان‌مند را نشان می‌دهد که شامل طیف گسترده‌ای از عناصر، از توسعه «محموله مخرب» تا «حرکت جانبی[۵]» در داخل یک محیط می‌گردد (Quintero-Bonilla, et. al, 2020: 14) زنجیره کشتن سایبری یک مفهوم نظامی است که به‌وسیله شرکت آمریکائی لاکهید مارتین توسعه یافت و به مراحلی اشاره می‌کند که مبتنی بر دکترین نظامی آمریکا و مفاهیم[۶] F2T2EA به هدف، دسترسی پیدا می‌کند. هفت مرحله تهاجم در زنجیره کشتن سایبری شامل شناسایی، ساخت سلاح، تحویل، بهره‌برداری نصب، فرماندهی و کنترل (C2) و اقدامات بر روی اهداف[۷] می‌باشد (Sánchez del Monte, et. al, 2023: 6).

چارچوب حملات مایتره

چارچوب حملات مایتره پایگاه دانشی رایگان است که اطلاعات جامع و برخطی را دررابطه‌با حملات سایبری برای سازمان‌ها در قالب یک ماتریس فراهم می‌آورد. سازمان غیرانتفاعی آمریکایی مایتره پایگاه دانشATT&CK را توسعه داده است. کلمه ATT&CK[۸] به معنی راهکنش‌ها[۹]، فن‌ها[۱۰] و دانش عمومی تهاجمی است که این پایگاه دانش را می‌سازند (Sánchez del Monte, et. al, 2023: 8)که شامل:

---

[1] Lockheed Martin’s Cyber Kill Chain
[2] MITRE ATT&CK framework
[3] Diamond Model
[4] Cyber Kill Chain
[5] lateral movement
[6] F2T2EA (Find, Fix, Track, Target, Engage, Assess)
[7] Reconnaissance, Weaponization, Delivery, Exploitation, Installation, Command and Control (C2) and Actions on objectives
[8] Adversarial Tactics, Techniques and Common knowledge
[9] Tactics
[10] Techniques

- راهکنش‌ها: اهداف تهاجمی راهکنشی در طول حمله (ستون‌های پایگاه دانش)
- فن‌ها: روش‌های حصول به اهداف تهاجمی راهکنشی (هر خانه ذیل ستون در پایگاه دانش)

راهکنش‌های تهاجمی سایبری به انواع زیر تقسیم‌بندی می‌شوند(Liu, et. al, 2020: 11):

۱- شناسایی[1]: گردآوری اطلاعاتی مثل اطلاعات مرتبط با سازمان هدف برای طرح‌ریزی اقدامات آفندی در آینده

۲- توسعه منبع:[2] ایجاد منابع مثل منطقه الکترونیک، خرید دامنه و استقرار زیرساخت‌های فرماندهی و واپایش برای پشتیبانی از عملیات

۳- دسترسی اولیه:[3] تلاش برای نفوذ به درون شبکه سازمان با روش‌هایی مثل فیشینگ

۴- اجرا[4]: تلاش برای اجرای بدافزارهایی همانند یک ابزار دسترسی از راه دور

۵- تداوم[5]: تلاش برای حذف رد پا با روش‌هایی مثل تغییر تنظیمات سامانه هدف، ساخت نام کاربری، احراز هویت و... .

۶- ارتقای دسترسی:[6] تلاش برای رسیدن به مجوزهای سطح بالاتر با روش‌هایی مثل کمک‌گرفتن از یک آسیب‌پذیری برای ارتقای دسترسی

۷- گریز از دفاع:[7] تلاش برای شناسایی‌نشدن با روش‌هایی مثل استفاده از پردازش‌های مورداعتماد برای مخفی‌کردن بدافزار

۸- دسترسی به اعتبارنامه‌ها:[8] سرقت نام حساب‌های کاربری و گذرواژه‌ها با روش‌هایی مثل کلیدخوان[9]

۹- کشف:[10] تلاش برای درک محیط مثل چیزهایی که مهاجم می‌تواند به واپایش خودش درآورد

---

[1] Reconnaissance
[2] Resource Development
[3] Initial Access
[4] Execution
[5] Persistence
[6] Privilege Escalation
[7] Defense Evasion
[8] Credential Access
[9] Keylogger
[10] Discovery

۱۰- حرکت جانبی: حرکت در عرض محیط مثل استفاده از اعتبارنامه‌ها برای نفوذ به سامانه‌های موازی

۱۱- گردآوری[1]: گردآوری اطلاعاتی در راستای هدف تخاصم مثل دسترسی به داده‌ها در فضای ابری

۱۲- فرماندهی و واپایش: ارتباط‌گرفتن با سامانه‌های تحت نفوذ برای کنترل‌کردن آن‌ها با روش‌هایی مثل شبیه‌سازی شدآمد طبیعی برای ارتباط با یک شبکه قربانی

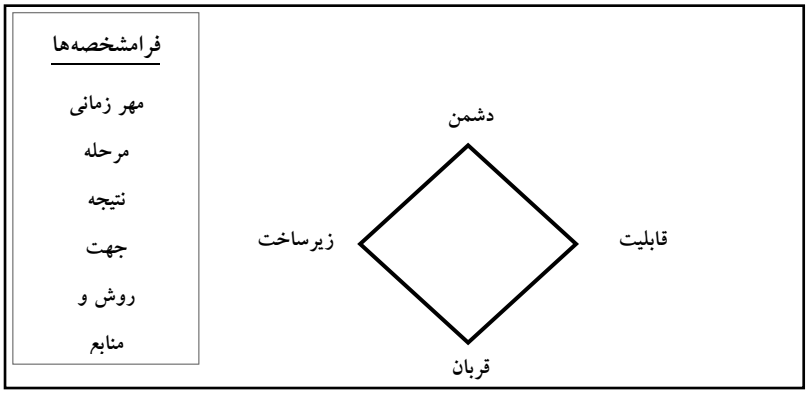
۱۳- دسترسی غیرمجاز به اطلاعات:[2] سرقت داده‌ها با روش‌هایی مثل انتقال داده‌ها به یک حساب ابری

۱۴- ضربه‌زدن:[3] دست‌کاری یا تداخل در داده‌ها و سامانه‌ها یا آسیب‌زدن به آن‌ها با روش‌هایی مثل رمزنگاری باج‌افزاری[4]

مدل الماس

مدل الماس مدلی است که اصول علمی را برای تجزیه‌وتحلیل نفوذ بررسی می‌کند و مهارت‌های مهاجم را به زیرساخت هدف مرتبط می‌کند. به عبارتی توانایی‌های تهاجمی درحال‌توسعه در زیرساختی که قربانی را هدف قرار می‌دهد مطالعه می‌کند. این مدل به‌عنوان مجموعه‌ای از رویدادها که چهار ویژگی کلیدی (دشمن، قابلیت، زیرساخت و قربانی[5]) را برای هر حمله بیان می‌کند، ساختاریافته است. این خصوصیات ارتباط نزدیکی با یکدیگر دارند و به‌عنوان رئوس یک لوزی الماس مطابق شکل (۱) پیکربندی شده‌اند (Zacchia, et. al, 2022: 18). علاوه بر چهار ویژگی کلیدی، فرا مشخصه‌ها[6] نیز تعریف شده‌اند که‌بخش عمده‌ای از مدل را تشکیل می‌دهند: (مهر زمانی، مرحله، نتیجه، جهت، روش، و منابع[7]) (Sánchez del Monte, et. al, 2023: 5).

[1] Collection
[2] Exfiltration
[3] Impact
[4] Ransomware
[5] Adversary, capability, infrastructure, and victim
[6] Meta-characteristics
[7] Time stamp, phase, outcome, direction, methodology, and resources



شکل (۱)مدل تجزیه‌وتحلیل نفوذ

زیرساخت نیز ساختارهای فیزیکی یا منطقی را توصیف می‌کند که برای به‌کارگیری قابلیت‌ها استفاده می‌شوند که از نظر نوع و حجم بسیار متغیر هستند. برای مثال نشان‌دهنده IP، دامنه‌ها، نشانی‌های منطقه الکترونیکی و دستگاه‌های فیزیکی هستند. دشمن شخص یا گروهی از افراد هستند که مسئول بهره‌برداری از قابلیت‌ها است. این قابلیت شامل ابزارهایی است که دشمن در بهره‌برداری از عملیات استفاده می‌کند: مانند آسیب‌پذیری‌های قربانی و مجموعه ابزارهایی که مهاجم می‌تواند از طریق آنها عملیات یک حمله را هدایت می‌کنند، درک می‌شود (Ibid: 19).

هوش مصنوعی

هوش مصنوعی به توانایی یک ماشین برای فکرکردن و انجام وظایف مانند یک انسان اشاره دارد. یادگیری ماشینی[1] زیرمجموعه‌ای از هوش مصنوعی است که نشان‌دهنده توانایی پیش‌بینی و تشخیص الگوی ماشین است. یادگیری ماشین، رایانه را برای انجام کارهای خاص، ارزیابی داده‌ها و استفاده از آن‌ها در تصمیم‌گیری آموزش می‌دهد. در عوض، الگوریتم‌های آن از یک‌چرخه بی‌پایان از ورودی داده‌ها و پیش‌بینی‌های خروجی پیروی می‌کنند. پیش‌بینی‌ها با داده‌های جدید بررسی می‌شوند و عامل‌های الگوریتم براین‌اساس بهبود می‌یابند (Alwasel, et. al, 2023: 1-4).

[1] Machine Learning

چارچوب نظری تحقیق

باتوجه‌به این واقعیت که ابزارهای هوش مصنوعی در حال حاضر درحال‌توسعه منبع‌باز[1] هستند، منطقی است که انتظار داشته باشیم که فناوری‌های هوش مصنوعی ممکن است برای ایجاد انواع جدیدی از عملیات سایبری پیشرفته و پیچیده استفاده شوند. در این‌بخش طیف وسیعی از کاربردهای عملی را نشان می‌دهیم که هوش مصنوعی می‌تواند در عملیات سایبری تهاجمی به کار رود. برخی از آنها در حال حاضر به شکل محدود در عمل استفاده می‌شوند، اما می‌توانند با پیشرفت فناوری بیشتر در آینده افزایش یا تقویت شوند (Thanh, et. al, 2019: 30). برای بررسی کاربردهای هوش مصنوعی و یادگیری ماشین در عملیات سایبری ما در این تحقیق چارچوب حملات مایتره را به این دلیل که؛ این مدل جدیدترین، گسترده‌ترین و ساختاریافته‌ترین مدلی است که مجموعه وسیعی از فرایندهای قابل‌شناسایی را ارائه می‌کند (Zurowski, et. al, 2022: 2)، مبنای مطالعه خود قرار دادیم:

به‌کارگیری هوش مصنوعی در عملیات سایبری تهاجمی

هوش مصنوعی در شناسائی و دسترسی

سن تزو[2] نقل می‌کند که اگر دشمن و خودتان را بشناسید، لازم نیست از نتیجه جنگ بترسید (Yamin, et. al, 2022: 1). توانایی پردازش کارآمد داده‌های متنی هوش مصنوعی با سرعت و دقتی فراتر از عامل انسانی آن را به ابزاری قدرتمند برای شناسایی تبدیل نموده است. محققان چهار کاربرد هوش مصنوعی برای گام اول یا همان «شناسایی» در عملیات سایبری را معرفی می‌کنند که عبارت‌اند از جمع‌آوری و تجزیه‌وتحلیل اطلاعات، نمایه‌سازی هدف[3]، تشخیص آسیب‌پذیری و پیش‌بینی نتیجه (Zouave, et. al, 2020: 17).

در جمع‌آوری اطلاعات از اهداف مبتنی بر وب[4] چون ارزیابی ابزارها، تاکتیک‌ها، پیکربندی، اسکن آسیب‌پذیری و... در هدف، فناوری غالب پردازش زبان طبیعی است و همه آن اطلاعات به همراه سایر اطلاعات تکمیلی چون نام سازمان، مکان، ساعات کار، مشخصات شبکه، کاربران شبکه و سازمان و... جمع‌آوری و در قالب وب معنائی برای

[1] open source
[2] Sun Tzu
[3] Target Profiling
[4] Web

ماشین قابل‌فهم می‌شوند (Ibid: 18). به‌عنوان‌مثال استفاده از روش‌های مختلف یادگیری عمیق از جمله شبکه عصبی تکراری[۱]، حافظه کوتاه‌مدت[۲] با قابلیت‌های تلفیقی[۳] و دیگر طبقه‌بندی‌کننده‌های یادگیری ماشین برای تجزیه‌وتحلیل گزارش‌های DNS در یک شبکه محلی[۴]، کشف نام دامنه و شناسایی فعالیت‌های آسیب‌پذیر مناسب است (Yamin, et. al, 2022: 21).

برای مدل‌سازی مفهومی داده‌های جمع‌آوری شده از چارچوب توصیف منابع[۵]، زبان هستی‌شناسی وب[۶] و زبان قوانین وب معنایی[۷] استفاده می‌گردد. از هستی‌شناسی وب برای استخراج ادعاهای قابل‌فهم ماشین نیز استفاده شده است. منطق نرم احتمالی[۸] همراه با مدل‌های پنهان مارکوف[۹] برای تشخیص روابط در داده‌های معنایی و پیش‌بینی حملات سایبری آینده بر اساس داده‌ها استفاده می‌شود (Zouave, et. al, 2020: 19).

در گام دوم از مرحله شناسائی به‌کارگیری الگوریتم‌های شبکه‌های عصبی پیچشی یا هم‌گشتی[۱۰] و هستی‌شناسی منطق فازی[۱۱] روش‌های غالب پردازش زبان طبیعی در «*نمایه‌سازی هدف*» هستند. (Ibid: 19). در گام سوم و تشخیص آسیب‌پذیری ادبیات شناسایی شده از طریق تحلیل متنی منابع آنلاین و آسیب‌پذیری‌های مرتبط با فعالیت‌ها در مرورگرهای وب به برخی فن‌های کاربردی هوش مصنوعی چون روش‌های رگرسیون و طبقه‌بندی چون طبقه‌بندی‌کننده ماشین بردار پشتیبان[۱۲] اشاره کرده‌اند. در برخی مطالعات نیز از الگوریتم‌های یادگیری ماشین چون SVM، جنگل تصادفی[۱۳]، بیز ساده[۱۴]، رگرسیون لجستیک[۱۵] برای سامانه امتیازدهی به آسیب‌پذیری‌ها[۱۶] بر اساس

[1] Recurrent neural network (RNN)
[2] long short-term memory (LSTM)
[3] Consolidated capabilities
[4] local area network (LAN)
[5] resource description framework (RDF)
[6] web ontology language (OWL)
[7] semantic web rule language (SWRL)
[8] Probabilistic soft logic (PSL)
[9] Hidden Markov models (cascading HMM)
[10] Convolutional neural networks (CNNs)
[11] fuzzy logic ontology
[12] Support Vector Machine (SVM)
[13] random forest (RF)
[14] naïve Bayes (NB)
[15] logistic regression (LOG-REG)
[16] Vulnerability Scoring System (CVSS)

معیارها و پیش‌بینی بهره‌برداری از آن‌ها در شرایط مختلف استفاده شده است در مواردی استخراج ویژگی‌های آسیب‌پذیری با استفاده از الگوریتم شبکه‌های عصبی پیش‌خور[1]، شبکه عصبی پیچشی و یک مدل طبقه‌بندی‌کننده انجام شده است. در گام چهارم از شناسائی و «*پیش‌بینی نتیجه*»، یادگیری ماشین می‌تواند در کنار روش‌های ارزیابی و شبیه‌سازی عملیات سایبری نتایج تحلیل شده‌ای را ارائه دهد (Ibid: 21).

هوش مصنوعی و اجرا

بدافزارها به طور خودکار بر اساس وضعیت سامانه میزبان و مبتنی بر آموزش‌های شرطی که قبلاً دیده است، سفارشی عمل می‌نماید. بدافزار هوشمند می‌تواند فن‌های حرکت جانبی را بسته به محیط خود انتخاب کند. به‌جای سوءاستفاده از آسیب‌پذیری‌ها، می‌تواند به حملات جستجوی فراگیر[2] سوئیچ کند یا حتی یک کلیدخوان[3] برای گرفتن اعتبار، دسترسی به اعتبارنامه‌ها و سرقت حساب‌های کاربری نصب کند. سپس بدافزار به طور خودکار موفق‌ترین اقدام را برای رسیدن به هدف و ارتقای دسترسی انتخاب می‌کند (Thanh, et. al, 2019: 29). در مطالعات انجام شده، الگوریتم ژنتیک[4] بر اساس الگوریتم کلونی زنبورهای عسل مصنوعی[5] و الگوریتم بازپخت شبیه‌سازی شده[6] برای تقویت تشخیص خودکار آسیب‌پذیری در سلاح سایبری و تولید سوءاستفاده بعدی معرفی شده است. در برخی سلاح‌های سایبری مولد از شبکه متخاصم مولد[7] در الگوریتم یادگیری عمیق، استفاده شده است. در سلاح سایبری حمله به رمز عبور، الگوریتم یادگیری ماشین منبع‌باز قابلیت حدس رمز عبور را بر اساس الگوهای موجود به سلاح سایبری می‌دهد (Zouave, et. al, 2020: 24). الگوریتم‌های GAN را می‌توان با یادگیری از پایگاه‌های داده رمز عبور فاش شده برای اعمال هوشمندانه گذرواژه‌ها استفاده کرد، محققان این رویکرد را با استفاده از الگوریتم RNN در فرایند تولید، بهبود بخشیده‌اند (Mirsky, et. al, 2023: 12). الگوریتم‌های متکی بر k-نزدیک‌ترین

[1] feed forward neural networks (FFN)
[2] brute-force attack
[3] key-logger
[4] genetic algorithm
[5] Artificial bee colony (ABC)
[6] Simulated Annealing (SA)
[7] Generative adversarial network (GAN)

همسایگان[۱]، رگرسیون لجستیک، طبقه‌بندی‌کننده بردار پشتیبان خطی[۲]، درخت تصمیم، جنگل تصادفی، درخت رگرسیون تقویت‌کننده گرادیان[۳]، ماشین بردار پشتیبان، و پرسپترون چندلایه(MLP) در سرقت داده‌های صفحه‌کلید قربانی توسط سلاح‌های سایبری کاربرد داشته‌اند. علاوه بر این طبقه‌بندی‌کننده‌هایی مانند درخت تصمیم، بیز ساده، پرسپترون چندلایه[۴] و ماشین بردار پشتیبانی برای پیش‌بینی قدرت رمز عبور استفاده شده است. در مواردی شبکه عصبی پیچشی و الگوریتم تطبیقی[۵] برای شکست‌دادن کپچاهای[۶] مبتنی بر متن با چند رقم در سلاح سایبری حمله به کپچا به‌کارگیری شده است. در تولید بدافزارهای خودآموز[۷]، از خوشه‌بندی k-means برای طبقه‌بندی داده‌های واپایش منطقی سامانه هدف و سوءاستفاده از قدرت شناسائی آسیب‌پذیری سامانه هدف استفاده شده و در این بدافزارها توزیع گاوسی[۸] برای شناسایی اثرات حمله بر سامانه هدف، جمع‌آوری داده و اجرای کدهای مخرب و اخلال در سامانه قربانی بکار گرفته شده است (Zouave, et. al, 2020: 24).

در مطالعات اخیر استفاده از الگوریتم CNN را دارای بالاترین کاربرد در تولید تسلیحات سایبری در حملات دسترسی و نفوذ معرفی نموده است و دیگر الگوریتم‌های GAN و RNN در مراتب بعد کاربرد قرار دارند. الگوریتم GAN در ایجاد قابلیت پنهان‌سازی هوشمند بدافزارهای متخاصم و URLهای بدافزار غیرقابل‌شناسایی بکار گرفته شده و از LSTM برای تولید بارهای مخرب فراری خودکار استفاده شده است (Guembe, et. al, 2022: 10). از زنجیره‌های مارکوف[۹] نیز برای ایجاد خودکار پیام‌های متنی در ابزارهای فیشینگ ("طعمه کلاهبرداری") استفاده می‌کنند (Zouave, et. al, 2020: 19).

---

[1] k-Nearest Neighbors (KNN)
[2] Support Vector Classifier
[3] gradient boosting regression tree(GB)
[4] multilayer perceptron (MLP)
[5] Adaptive Algorithm (AA)
[6] Captcha
[7] Self-Learning Malware
[8] Gaussian distribution (GD)
[9] Markov chain

هوش مصنوعی و گمنامی

در مرحله «گریز از دفاع» توسط سامانه‌های دفاعی قربانی، دودسته از بدافزارها به نام‌های بدافزار مخفی ساز هوشمند[1] و بدافزار فراری[2] شناسائی شده‌اند یافته‌های تحقیقات اخیر حاکی از این است که الگوریتم GAN بیشترین کاربرد را در مرحله تحویل از زنجیره کشتن سایبری دارد درحالی‌که الگوریتم‌های DNN و LSTM در جایگاه بعدی قرار دارند (Guembe, et. al, 2022: 14).

الگوریتم‌های GAN می‌توانند برای پنهان‌کردن قصد بدافزار از یک تحلیل‌گر نیز استفاده شوند. در این پنهان‌سازی استفاده مجدد از بدافزار فعال نگه داشته شده، اهداف و زیرساخت مهاجم پنهان شده و مدت حمله طولانی می‌گردد. به این معنا که یک نرم‌افزار موجود را برداشته و قطعه دیگری را که از نظر عملکردی معادل و مشابه آن است (مشابه ترجمه در NLP) جایگزین می‌گردد. به‌عنوان‌مثال، DeepObfusCode از شبکه‌های عصبی بازگشتی برای تولید کد رمزگذاری شده استفاده می‌کند. روش دیگر، درهای پشتی[3] را می‌توان در پروژه‌های منبع‌باز[4] کاشت و با روش‌های مشابه پنهان کرد (Mirsky, et. al, 2023: 10).

هوش مصنوعی و اقدام

اقدام و ضربه‌زدن شامل، تاکتیک‌هایی است که برای برهم‌زدن دردسترس‌بودن یا به تغییر یکپارچگی با دست‌کاری فرایندهای تجاری و عملیاتی استفاده می‌شود. استفاده از قابلیت بهره‌برداری خودکار از آسیب‌پذیری‌ها در عین مخفی‌سازی هوشمند، در تولید کد حمله خودکار از اولویت بالائی برخوردار است (Zouave, et. al, 2020: 24). ازاین‌رو باتوجه‌به اینکه هوش مصنوعی داده‌ها و پروتکل‌ها را سریع‌تر، پیوسته‌تر و با تداوم بیشتری نسبت به انسان‌ها شناسایی و بهره‌برداری می‌کند، به‌عنوان ابزاری برای بهره‌برداری از آسیب‌پذیری‌ها در راهکنش ضربه، نقش‌برجسته‌ای دارد (Ibid: 35). الگوریتم‌های شناسایی شده در مراحل مختلف چارچوب حملات مایتره به طور خلاصه در شکل ۲ نمایش داده شده‌اند:

[1] Intelligent concealment malware
[2] Evasive malware
[3] Backdoors
[4] Open source projects

ترکیبی	دسترسی به اطلاعات	فرماندهی و کنترل	گردآوری	حرکت جانبی	کشف	دسترسی به اعتبارنامه	گزیز	ارتقا دسترسی	تداوم	اجرا	دسترسی اولیه	توسعه منبع	شناسایی
GD	GD	K-means clustering	K-means clustering	ABC	RNN	ABC	GAN	SVM	Torch RNN	K-means clustering	HMM	SVM	RNN
CNNs	CNN		GD	SA	LSTM	SA	DNN	RF	KNN	GD	SVM	NB	LSTM
GAN				GAN	RDF	GAN	LSTM	NB	SVC	CNN	NB	MLP	RDF
RNN				Torch RNN	OWL	KNN		LOG-REG	GB		MLP		OWL
LSTM				KNN	SWRL	SVC		CVSS	MLP				SWRL
				SVC	PSL	GB		FNN					PSL
				GB	HMM	MLP		CNN					HMM
				MLP	CNNs			ABC					CNNs
					FLO			SA					FLO
								GAN					

شکل (۲): الگوریتم‌های هوش مصنوعی در چارچوب حملات مایتره

روش‌شناسی تحقیق

این پژوهش از حیث جهت‌گیری، از نوع مطالعات کاربردی توسعه‌ای به شمار می‌رود و از نظر ماهیت داده‌ها، از نوع پژوهش‌های کیفی و از نظر قطعیت داده‌ها، از نوع اکتشافی است. روش تجزیه‌وتحلیل اطلاعات، تحلیل مضمون متنی است. تحلیل مضمون، روشی برای شناخت، تحلیل و گزارش الگوها (مضامین) موجود در داده‌های کیفی است. این روش، فرایندی برای تحلیل داده‌های متنی است و داده‌های پراکنده و متنوع را به داده‌هایی غنی و تفصیلی تبدیل می‌کند (عابدی جعفری و دیگران، ۱۳۹۰: ۱۵۳). مضامین از درون کدها استخراج می‌شوند و با تحلیل تفسیری داده‌ها توسعه می‌یابند؛ مضامین فراگیر، هسته و کانون شبکه‌ای مضمونی را تشکیل می‌دهد. (همان: ۱۵۴). محققین در این مقاله، باتوجه‌به هدف تحقیق، داده‌ها را برای شناسایی مضامین مطالعه نموده و در ادامه، مضامین پایه، سازمان‌دهنده و فراگیر را تشخیص می‌دهند. برای سنجش اعتبار پژوهش از آزمون اعتبار محتوا استفاده شده است. اعتبار به این نکته اشاره دارد که؛ داده‌های به‌دست‌آمده تا چه میزانی، معنای واقعی مفهوم مورد بررسی را به‌اندازه کافی منعکس می‌نماید (ببی، ۱۳۹۲: ۳۳۵). در واقع اعتبارسنجی به این مفهوم است که، «آیا واقعاً همان چیزی را می‌سنجیم که قصد سنجش آن را داریم» (بیکر، ۱۳۹۲: ۱۳۸). لذا در پژوهش پیشرو به‌منظور اعتباربخشی به یافته‌ها از سه راهبرد استفاده شده است. نخست، بازبینی به وسیله همکار پژوهشی؛ در این مرحله یافته‌های پژوهش در اختیار گروه راهنما قرار گرفت و پس از بازبینی و انجام اصلاحات،

بازبینی در کدها، تِم‌های فرعی اولیه و ثانویه انجام شد. در مرحله دوم، از راهبرد بازبینی همتایان استفاده شد. در این مرحله یافته‌ها در اختیار محققان بی‌طرف خارج از گروه پژوهشی قرار گرفت که تجربه انجام پژوهش‌های کیفی را دارند. با اعمال نظرات، اعتبار داده‌ها افزایش یافت. علاوه بر دو راهبرد فوق، از قضاوت و نظر ده نفر از پژوهشگران حوزه پژوهش موردنظر متشکل از اعضای هیئت‌علمی و پژوهشگران دانشگاه عالی دفاع ملی در قالب پنل «دلفی» استفاده شده است. همچنین برای سنجش پایایی در این پژوهش، از شاخص درصد توافق (ضریب کاپای کوهن) استفاده شده است. به‌این‌ترتیب، پس از پایان کدگذاری، حدود ۲۰ درصد از کدگذاری‌ها انتخاب و بار دیگر از سوی پژوهشگران کدگذاری شد. میزان تشابه فراوانی کدها با استفاده از این شاخص محاسبه شد که برای همه مؤلفه‌ها بالاتر از ۷۰ درصد تعیین شد.

یافته‌های پژوهش

در این پژوهش، برای استخراج شاخص‌ها و مضامین مرتبط، به بررسی هدفمند منابع پرداخته شده است. در این راستا، در جستجوی اولیه در گوگل، گوگل‌اسکولار و پایگاه‌های علمی معتبر (ساینس دایرکت، اسپرینگر، امرالد، تیلور و فرانسیس) با جستجوی واژه‌های هوش مصنوعی و عملیات سایبری تهاجمی در مجموع، ۱۲۱ نتیجه یافت شد. در مرحله اول، ۴۸ منبع به دلیل اینکه در ارتباط مستقیم با موضوع نبودند، حذف شدند؛ در مرحله دوم، ۱۷ منبع تکراری از فهرست منابع کنار گذاشته شدند و در مرحله پایانی، برای استخراج شبکه مضامین، ۵۶ منبع موردمطالعه دقیق قرار گرفت که مضامین و کدها از ۲۷ منبع استخراج گردید و از ۲۹ منبع دیگر، به دلیل هم‌پوشانی با سایر منابع، مضامینی برای استخراج یافت نشد. بر این اساس نهایتاً ۲۷ منبع به شرح جدول (۱) برای مطالعهٔ عمیق گزینش شدند.

جدول(۱) تعداد مقالات منتخب از پایگاه‌های علمی

نام پایگاه علمی	ساینس دایرکت	اسپرینگر	امرالد	تیلور و فرانسیس	آرکسیو	مگیران	مجموع
تعداد	۱۰	۴	۶	۲	۳	۲	۲۷

در ادامه نمونه مضامین پایه استخراج شده از متون مذکور، در قالب جدول (۲) نمایش‌داده‌شده است.

جدول(۲) مضامین استخراج شده از الگوریتم‌های هوش مصنوعی و عملیات سایبری

ردیف	گویه‌های متن	منبع	کدهای استخراج شده
۱	استفاده از روش‌های مختلف یادگیری عمیق از جمله شبکه عصبی تکراری(RNN) ، حافظه کوتاه‌مدت(LSTM) با قابلیت‌های تلفیقی و دیگر طبقه‌بندی‌کننده‌های یادگیری ماشین برای تجزیه‌وتحلیل گزارش‌های DNS در یک شبکه محلی، کشف نام دامنه و شناسایی فعالیت‌های آسیب‌پذیر مناسب است	Yamin, et. al, 2022: 21	جمع‌آوری اطلاعات
۲	برای مدل‌سازی مفهومی داده‌های جمع‌آوری شده از چارچوب توصیف منابع (RDF)، زبان هستی‌شناسی وب (OWL) و زبان قوانین وب معنایی (SWRL) استفاده می‌گردد. از هستی‌شناسی وب (OWL) برای استخراج ادعاهای قابل‌فهم ماشین نیز استفاده شده است. منطق نرم احتمالی (PSL) همراه با مدل‌های پنهان مارکوف (HMM آبشاری) برای تشخیص روابط در داده‌های معنایی و پیش‌بینی حملات سایبری آینده بر اساس داده‌ها استفاده می‌شود	Zouave, et. al, 2020: 19	مدل‌سازی درک اطلاعات و پیش‌بینی
۳	الگوریتم‌های شبکه‌های عصبی پیچشی یا هم‌گشتی (CNNs) و هستی‌شناسی منطق فازی (FLO) روش‌های غالب پردازش زبان طبیعی در «نمایه‌سازی هدف» هستند	Zouave, et. al, 2020: 19	نمایه‌سازی هدف
۴	تشخیص آسیب‌پذیری از طریق تحلیل متنی منابع آنلاین و آسیب‌پذیری‌های مرتبط با فعالیت‌ها در مرورگرهای وب به برخی فن‌های کاربردی هوش مصنوعی چون روش‌های رگرسیون و طبقه‌بندی چون طبقه‌بندی‌کننده ماشین بردار پشتیبان (SVM) اشاره کرده‌اند. در برخی مطالعات نیز از الگوریتم‌های یادگیری ماشین چون SVM، جنگل تصادفی (RF)، بیز ساده (NB)، رگرسیون لجستیک (LOG-REG) برای سامانه امتیازدهی به آسیب‌پذیری‌ها بر اساس معیارها (CVSS) و پیش‌بینی بهره‌برداری از آن‌ها در شرایط مختلف	Zouave, et. al, 2020: 21	ارتقای دسترسی از طریق شناسائی آسیب‌پذیری‌ها

ردیف	گویه‌های متن	منبع	کدهای استخراج شده
	استفاده شده است در مواردی استخراج ویژگی‌های آسیب‌پذیری با استفاده از الگوریتم شبکه‌های عصبی پیش‌خور (FNN)، شبکه عصبی پیچشی (CNN) و یک مدل طبقه‌بندی‌کننده انجام شده است		
۵	در مطالعات انجام شده، الگوریتم ژنتیک بر اساس الگوریتم کلونی زنبورهای عسل مصنوعی (ABC) و الگوریتم بازپخت شبیه‌سازی شده (SA) برای تقویت تشخیص خودکار آسیب‌پذیری در سلاح سایبری و تولید سوءاستفاده بعدی معرفی شده است. در برخی سلاح‌های سایبری مولد از شبکه متخاصم مولد (GAN) در الگوریتم یادگیری عمیق، استفاده شده است.	Zouave, et. al, 2020: 24	
۶	در سلاح سایبری حمله به رمز عبور، الگوریتم یادگیری ماشین منبع‌باز (Torch RNN) قابلیت حدس رمز عبور را بر اساس الگوهای موجود به سلاح سایبری می‌دهد و الگوریتم‌های متکی بر k-نزدیک‌ترین همسایگان (KNN)، رگرسیون لجستیک، طبقه‌بندی‌کننده بردار پشتیبان خطی (SVC)، درخت تصمیم (DT)، جنگل تصادفی (RF)، درخت رگرسیون تقویت‌کننده گرادیان (GB)، ماشین بردار پشتیبان (SVM)، و پرسپترون چندلایه (MLP) در سرقت داده‌های صفحه‌کلید قربانی توسط سلاح‌های سایبری کاربرد داشته‌اند.	Zouave, et. al, 2020: 24	حمله به رمز عبور (حدس، پیش‌بینی)، کلیدخوان
۷	علاوه بر این طبقه‌بندی‌کننده‌هایی مانند درخت تصمیم، بیز ساده (NB)، پرسپترون چند لایه (MLP) و ماشین بردار پشتیبانی (SVM) برای پیش‌بینی قدرت رمز عبور استفاده شده است. در مواردی شبکه عصبی پیچشی (CNN) و الگوریتم تطبیقی (AA)برای شکست‌دادن کپچاهای مبتنی بر متن با چند رقم در سلاح سایبری حمله به کپچا به‌کارگیری شده است. در تولید بدافزارهای خودآموز (SLM)، از خوشه‌بندی K-means برای	Zouave, et. al, 2020: 24	حمله به رمز عبور (پیش‌بینی)، حمله به کپچا، بدافزارها، کد مخرب و اخلال

ردیف	گویه‌های متن	منبع	کدهای استخراج شده
	طبقه‌بندی داده‌های واپایش منطقی سامانه هدف و سوءاستفاده از قدرت شناسائی آسیب‌پذیری سامانه هدف استفاده شده و در این بدافزارها توزیع گاوسی (GD) برای شناسایی اثرات حمله بر سامانه هدف، جمع‌آوری داده و اجرای کدهای مخرب و اخلال در سامانه قربانی بکار گرفته شده است		
۸	از الگوریتم CNNرا دارای بالاترین کاربرد در تولید تسلیحات سایبری در حملات دسترسی و نفوذ معرفی نموده است و دیگر الگوریتم‌های GANو RNNدر مراتب بعد کاربرد قرار دارند و از LSTM برای تولید بارهای مخرب فراری خودکار استفاده شده است	Guembe, et. al, 2022: 10	حملات دسترسی و نفوذ بارهای مخرب
۹	از زنجیره‌های مارکوف نیز برای ایجاد خودکار پیام‌های متنی در ابزارهای فیشینگ ("طعمه کلاهبرداری") استفاده می‌کنند.	Zouave, et. al, 2020: 19	بدافزار
۱۰	الگوریتم GANدر ایجاد قابلیت پنهان‌سازی هوشمند بدافزارهای متخاصم و URLهای بدافزار غیرقابل‌شناسایی بکار گرفته شده	Guembe, et. al, 2022: 10	پنهان‌سازی
۱۱	دودسته از بدافزارها به نام‌های بدافزار مخفی ساز هوشمند و بدافزار فراری شناسائی شده‌اند یافته‌های تحقیقات اخیر حاکی از این است که الگوریتم GAN بیشترین کاربرد را در مرحله تحویل از زنجیره کشتن سایبری دارد درحالی‌که الگوریتم‌های DNN و LSTM در جایگاه بعدی قرار دارند	Guembe, et. al, 2022: 14	

پس از پایان کدگذاری اولیه و احصاء ۳۹ کد، کدهای مشترک کنار یکدیگر قرار داده شدند تا درباره نحوه تلفیق و ترکیب کدهای مختلف برای تشکیل مضامین پایه تصمیم‌گیری شود.

در این مرحله مضامین پایه و کدها به کمک نفر از خبرگان (۵ نفر از اساتید و ۵ نفر از دانشجویان دانشگاه عالی دفاع ملی)مورد بررسی قرار گرفت تا درنهایت، جمع‌بندی نهایی حاصل شد. در گام بعدی، ماهیت و مفهوم هر مضمون پایه، به طور دقیق‌تر شناسایی و نام‌های مناسب برای مضمون‌ها انتخاب شد. سپس مضامین سطوح بالاتر

(مضامین سازمان‌دهنده) شناسایی شدند. بدین منظور، ابتدا مضامین اصلی که وجه اشتراک بسیاری با هم داشته و یا حول یک موضوع خاص بودند با هم ترکیب شده و یک مضمون سازمان‌دهنده را تشکیل دادند. مضامین سازمان‌دهنده، درواقع، مضامین اصلی مشابه را گروه‌بندی می‌کنند، سپس مضامین سازمان‌دهنده نیز بدین شیوه دسته‌بندی و ذیل مضامین فراگیر قرار گرفتند. نتایج این تحلیل، شامل شناسایی مضمون پایه، مضمون سازمان‌دهنده و مضمون فراگیر در جدول(۳) ارائه شده است.

جدول(۳) نتایج جمع‌بندی تحلیل مضمون متن

مضامین فراگیر	مضامین سازمان‌دهنده	مضامین پایه
به‌کارگیری الگوریتم‌های هوش مصنوعی در عملیات سایبری تهاجمی	شناسائی	جمع‌آوری اطلاعات
		مدل‌سازی و درک اطلاعات
		پیش‌بینی
		نمایه‌سازی هدف
		شناسائی آسیب‌پذیری‌ها
	اجرا	حمله به رمز عبور (حدس، سرقت)
		حمله به کپچا
		کد مخرب و خلال
		حمله به سطوح دسترسی
		نفوذ
		بدافزارها
		کلیدخوان
	گمنامی	مخفی‌سازی
		پنهان‌سازی
		فرار
	ضربه و اقدام	دست‌کاری اطلاعات
		پاک‌کردن اطلاعات
		رمزنگاری اطلاعات

شکل(۳) الگوی مفهومی به کارگیری الگوریتم‌های هوش مصنوعی در عملیات سایبری

نتیجه‌گیری

در این پژوهش با رویکرد تحلیل مضمون متنی، ۵۶ سند و پژوهش معتبر در حوزه موردمطالعه برای استخراج مضامین و کدها، موردمطالعه دقیق قرار گرفت.

بـر پایـهٔ یافته‌های پژوهـش، الگوی مفهومی به‌کارگیری الگوریتم‌های فناوری هوش مصنوعی در عملیات سایبری تهاجمی، شامل ۴ بعد شناسایی، اجراء، گمنامی و ضربه و اقدام است که متشکل از ۱۸ مؤلفه است. البته اعتبارسنجی دسته‌بندی‌ها، با مراجعه به جامعه خبرگی صورت گرفت که این مهم به تحصیل اعتبار خروجی نهایی پژوهش منجر شد.

باتوجه به اینکه این مطالعه، از نخستین مطالعات داخلی درباره الگوی به‌کارگیری هوش مصنوعی در عملیات سایبری است، بنابراین کمک شایان و درخور توجهی به ترویج گفتمان آن در جامعه، به‌ویژه جامعه علمی و اجرایی، خواهد کرد. نتایج فوق می‌تواند، بینشی ارزشمند به سیاست‌گذاران و تنظیم گران این حوزه که درصدد بهبود و ارتقا

توانمندی سایبری به واسطه به‌کارگیری هوش مصنوعی هستند، ارائه دهد. این پژوهش علاوه بر این، می‌تواند در سطوح عملی‌تر توسط مدیران سطوح میانی مورداستفاده قرار گیرد. در واقع نتایج این پژوهش به پژوهشگران و دست‌اندرکاران این حوزه کمک می‌کند تا بدانند، باید به چه ابعاد و مضامینی توجه کنند.

اساساً مهاجمان سایبری اعم از بازیگران دولتی و یا غیردولتی، به طور مداوم در حال تغییر و بهبود کارایی حملات خود هستند و بر استفاده از الگوریتم‌های هوش مصنوعی در فرایند حمله تأکید می‌کنند. این مطالعه قابلیت‌های تهاجمی هوش مصنوعی را که به مهاجمان اجازه می‌دهد حملاتی را در مقیاس بزرگ‌تر، با دامنه وسیع‌تر، سریع‌تر و دقیق‌تر آغاز کنند، شناسائی و معرفی می‌کند. این پژوهش نشان داد که ضمن فراگیر بودن کاربرد الگوریتم‌های هوش مصنوعی در تمامی مراحل ۱۴گانه چارچوب حملات مایتره، بیشترین الگوریتم‌ها، در مراحل شناسایی، ارتقای دسترسی، کشف، حرکت جانبی و دسترسی به اعتبارنامه به‌کارگیری شده‌اند.

همچنین باتوجه‌به نتایج حاصل از مصاحبه‌ها، علی‌رغم اینکه شرکت‌کنندگان معتقد بودند که در شناخت الگوریتم‌های هوش مصنوعی در عملیات سایبری و اشتراک‌گذاری اطلاعات در این زمینه، پیچیدگی‌های خاصی وجود دارد تأیید گردید که استفاده از الگوریتم‌های هوش مصنوعی می‌تواند مزایای مختلفی چون موارد زیر را به همراه داشته باشد:

- افزایش دقت و کارایی عملیات سایبری
- کاهش هزینه و زمان عملیات سایبری
- دستیابی به اهداف پیچیده‌تر

بااین‌حال، استفاده از الگوریتم‌های هوش مصنوعی در عملیات سایبری چالش‌هایی نیز به همراه دارد که نیاز است در استفاده از آن‌ها با آگاهی و احتیاط عمل کرد. باید بتوان اطمینان داشت که الگوریتم‌های هوش‌مصنوعی در شرایط مختلف عملکرد مناسبی خواهند داشت و نتایج قابل‌پیش‌بینی ارائه خواهند کرد. آن‌ها بایستی شفاف بوده و اهداف را به طور دقیق شناسایی کنند، حملات را به طور مؤثر اجرا کنند و به طور مسئولانه از خطاهای محاسباتی جلوگیری کنند.

در این تحقیق تلاش گردید تا با شناسائی و معرفی الگوریتم‌های توانمندساز در عملیات سایبری، ضمن ارائه الگوی مفهومی در راستای ارتقای توان رزم سایبری، زمینه علمی و دانش لازم جهت بازدارندگی سایبری نیز فراهم گردد.

پیشنهادها

با عنایت به کاربردی‌بودن این پژوهش پیشنهاد می‌گردد که:

۱- اتخاذ رویکرد دفاع فعالانه با بکارگیری الگوی مفهومی کاربست الگوریتم‌های هوش مصنوعی در عملیات سایبری به جای دفاع منفعلانه حاکم بر فضای سایبر توسط دستگاه‌های ذیربط

۲- ابزارهای متداول عملیات سایبری باقابلیت استفاده از الگوریتم‌های هوش مصنوعی شناسایی و در اولویت برنامه‌های توسعه و تجهیز سازمان‌ها قرار گیرد.

۳- در تحقیقات بعدی، مدل عملیاتی و اندازه گیری این الگو، با تدوین سنجه‌های خرد به همراه وزن دهی آنها انجام شود.

۴- با معرفی بیشتر هر یک از الگوریتم‌های شناسایی شده، زمینه آشنایی و استفاده بهتر از آن‌ها در عملیات سایبری فراهم گردد.

۵- در تحقیقات بعدی با استفاده از یافته‌ها و نتایج این تحقیق، الگوریتم‌های هوش مصنوعی در دفاع سایبری مطالعه و تبیین گردد.

قدردانی

از کلیهٔ اساتیدی که در فرآیند این تحقیق ما را با کمک‌های علمی خود یاری رساندند، صمیمانه سپاس‌گزاریم.

منابع

- ببی، ارل (۱۳۹۲)، *روش‌های تحقیق در علوم اجتماعی*: جلد نخست. ترجمه رضا فاضل. تهران: سازمان مطالعه و تدوین کتب علوم انسانی دانشگاه‌ها (سمت)، چاپ نهم.
- بیکر، ترز ال (۱۳۹۲). نحوهٔ انجام تحقیقات اجتماعی. ترجمه هوشنگ نایبی. تهران: نشر نی.
- رمضان‌زاده، مجید و همکاران (۱۳۹۹)، ارائه الگوی مفهومی ارزیابی قدرت سایبری نیروهای مسلح با تأکید بر بعد بازدارندگی، *فصلنامه مدیریت نظامی*، ۲۰(۷۸): ۹۲-۶۱.

- عابدی جعفری، حسن؛ و دیگران (۱۳۹۰)، تحلیل مضمون و شبکه مضامین: روشی ساده و کارآمد برای تبیین الگوهای موجود در داده‌های کیفی، *فصلنامه اندیشه مدیریت راهبردی*، سال پنجم، شماره ۲.
- فلیک، اووه (۱۳۹۳). *درآمدی بر تحقیق کیفی*. ترجمه هادی جلیلی. تهران: نشر نی.
- قاسمی، علی و همکاران (۱۴۰۲)، الگوی ارزیابی قدرت آفند سایبری ارتش ج. ا. ایران، *فصلنامه علوم و فنون نظامی*، ۱۹(۶۴): ۶۱-۳۶.
- Alwasel, A. , Mishra, S. , & AlShehri, M. , (2023), Attacks on The Vehicle Ad-Hoc Network from Cyberspace. *International Journal of Computing and Digital Systems.*
- Berzins, V. , (2022), Hybrid warfare: weaponized migration on the eastern border of the EU? *The Interdisciplinary Journal of International Studies*, 12(1), 19-19.
- FireEye, 2023, Highly Evasive Attacker Leverages SolarWinds Supply Chain to Compromise Multiple Global Victims With SUNBURST Backdoor. Available at: https: //www. fireeye. com/blog/threat-research/2020/12/evasive-attacker-leveragessolarwinds- supply-chain-compromises-with-sunburst-backdoor. html
- Guembe, B. , Azeta, A. , Misra, S. , Osamor, V. C. , Fernandez-Sanz, L. , & Pospelova, V. , (2022), *The emerging threat of ai-driven cyber-attacks: A review. Applied Artificial Intelligence.*
- Huskaj, G. , (2019), *The current state of research in offensive cyberspace operations.* In 18th European Conference on Cyber Warfare and Security (ECCWS 2019), 4-5 July 2019, Coimbra, Portugal (pp. 660-667). Academic Conferences and Publishing International Limited.
- Jp3-12, (2018), *Cyberspace Operations*, Joint Chiefs of Staff.
- Kaur, D. , Uslu, S. , & Durresi, A. , (2021), Requirements for trustworthy artificial intelligence–a review, In Advances in Networked-Based Information Systems: The 23rd International Conference on Network-Based Information Systems, Springer *International Publishing*.
- Laato, S. , Tiainen, M. , Najmul Islam, A. K. M. , & Mäntymäki, M. , (2022), How to explain AI systems to end users: a systematic literature review and research agenda. *Internet Research*.
- LeCun, Y. , (2022), A path towards autonomous machine intelligence, *Courant Institute of Mathematical Sciences*, New York University.
- Liu, C. , Singhal, A. , & Wijesekera, D. , (2020), Forensic analysis of advanced persistent threat attacks in cloud environments, In Advances in Digital Forensics XVI: 16th IFIP WG 11. 9 International **Conference**, New Delhi, India, January 6–8, 2020, Revised Selected Papers 16, Springer *International Publishing*.
- Maeda, R. , & Mimura, M. , (2021), *Automating post-exploitation with deep reinforcement learning. Computers & Security*.

- Mirsky, Y. , Demontis, A. , Kotak, J. , Shankar, R. , Gelei, D. , Yang, L. ,... & Biggio, B. , (2023), *The threat of offensive ai to organizations, Computers & Security*.
- Nguyen, H. V. , & Uehara, T. , (2023), Multilayer Action Representation based on MITRE ATT&CK for Automated Penetration Testing. *Journal of Information Processing*.
- Nichols, N. , (2023), The Future of Cyber Systems: Human-AI Reinforcement Learning with Adversarial Robustness, *In The Second Workshop on New Frontiers in Adversarial Machine Learning.*
- *Pecoraro,* G. , D'Amico, M. , & Romano, S. P. , (2021), *RedHerd: Offensive Cyberspace Operations as a Service. Signals.*
- Quintero-Bonilla, S. , & Martín del Rey, A. , (2020), *A new proposal on the advanced persistent threat: A survey. Applied Sciences.*
- Salik, H. , (2022), *Offensive Cyber Operations: Failure to Dissuade Nation-State Adversaries in Cyberspace* (Doctoral dissertation, University of the Cumberlands).
- Sánchez del Monte, A. , & Hernández-Álvarez, L. , (2023), *Analysis of Cyber-Intelligence Frameworks for AI Data Processing. Applied Sciences*, 13(16), 9328.
- Thanh, C. T. , & Zelinka, I. , (2019), *A survey on artificial intelligence in malware as next-generation threats. In Mendel* (Vol. 25, No. 2).
- Thiebes, S. , Lins, S. , & Sunyaev, A. , (2021), *Trustworthy artificial intelligence, Electronic Markets.*
- Walker-Munro, B. , Mount, D. , & Ioannou, R. , (2023), *The Hacker Strikes Back: Examining the Lawfulness of "Offensive Cyber" Under the Laws of Australia.*
- Westbrook, T. , (2023), A Taxonomy of Radiofrequency Jamming and Spoofing Strategies and Criminal Motives. *Journal of Strategic Security*.
- Yamin, M. M. , Ullah, M. , Ullah, H. , Katt, B. , Hijji, M. , & Muhammad, K. , (2022), *Mapping Tools for Open-Source Intelligence with Cyber Kill Chain for Adversarial Aware Security*. **Mathematics**.
- Zacchia, N. , Dondo, M. , & Taylor, A. , (2022), *Concepts for novel capabilities supporting cyber threat intelligence.*
- Zouave, E. , Bruce, M. , Colde, K. , Jaitner, M. , Rodhe, I. , & Gustafsson, T. , (2020), *Artificially intelligent cyberattacks*, Swedish Defence Research Agency, **FOI**.
- Zurowski, S. , Lord, G. , & Baggili, I. , (2022), A quantitative analysis of offensive cyber operation (OCO) automation tools. *In Proceedings of the 17th International Conference on Availability, Reliability and Security.*