Cyber Deception powered by Artificial Intelligence: Overview, Gaps, and Opportunities

Pedro Beltrán López
Department of Information and Communications Engineering
University of Murcia
pedro.beltranl@um.es

Pantaleone Nespoli
Department of Information and Communications Engineering
University of Murcia
pantaleone.nespoli@um.es

Manuel Gil Pérez
Department of Information and Communications Engineering
University of Murcia
mgilperez@um.es

***Abstract*—Cybersecurity is developing rapidly, and new methods of defense against attackers are appearing, such as Cyber Deception (CYDEC). CYDEC consists of deceiving the enemy who performs actions without realizing that he/she is being deceived. In this article, we conduct a comprehensive review of the CYDEC paradigm, addressing its main techniques and tools to its most relevant applications and highlighting the principal benefit in each scenario. Furthermore, we highlight the potentially revolutionary use of Artificial Intelligence (AI) in conjunction with CYDEC, analyzing the most significant gaps in the field of CYDEC-AI and, at the same time, identifying the most promising opportunities to address these gaps and present to the scientific community the next steps in the field of CYDEC-AI. Thanks to the research that follows from this paper, we will better understand the great potential for the use of CYDEC-AI in modern defense systems.**

***Index Terms*—Cyber Deception, Artificial Intelligence, Cybersecurity, Cyber Defense**

I. Introduction

In an increasingly connected and digitized world [1], the proliferation of users, devices, and applications has become overwhelming. This technology boom has been accompanied by an avalanche of data [2], which is of critical value and subject to strict privacy and confidentiality regulations. In this scenario, the security and reliability of systems become essential and non-negotiable in any sector of society and industry, especially when it comes to critical infrastructures where ensuring security is a top priority.

The cybersecurity landscape is exacerbated by the continued increase in the number and complexity of cyber attacks [3]. Malicious actors have developed increasingly sophisticated techniques to infiltrate systems, steal sensitive information, alter records, and, in most cases, cause significant damage. The expansion of threats such as ransomware [4], phishing, and cyber espionage constantly threatens individuals and organizations. This huge increase in cyber attacks is due to the economic benefit it brings to cybercriminals, thus creating an increasingly booming data market surpassing different countries in terms of Gross Domestic Product (GDP) [5].

Another critical arena that is being targeted by cybercriminals is the military environment, whose cyberspace has become a new battleground where nations measure their forces in a scenario that transcends physical borders [6]. In this sense, cyber capabilities have become crucial to many nations' national defense and security strategies as conflicts and tensions increasingly shift to the digital realm. States seek not only to protect their critical infrastructure and national secrets but also to use cyberspace to gather intelligence, conduct covert operations, and, in some cases, exert influence on the international stage [7].

In this context, a solid cybersecurity strategy is essential to protect against these constantly evolving threats. This strategy focuses not only on detection and response to incidents but also on proactive prevention or active defense [8]. This cybersecurity strategy seeks to fortify systems and safeguard organization's assets. It also allows for the anticipation and blocking of attacks intended to disable or disrupt entity's normal operations, which could result in catastrophic consequences, both financial and operational [9].

Due to these needs, new concepts and novel techniques are appearing to which special attention should be paid to defend against cyber attacks successfully. Specifically, one that has caught the attention of the research community and industry is Cyber Deception (CYDEC) [10], or Cyber Military Deception (CYMILDEC) [11] as a counterpart to CYDEC in military environments. In particular, CYDEC is defined as a cybersecurity strategy that involves deliberately implementing some form of deception in a environment to confuse and deter cyber attackers. Figure 1 shows the main advantages and use cases of CYDEC. These use cases are divided into three: slowing attacker, learning and protect assets, and each of them has several advantages that enhance the effectiveness of CYDEC techniques to provide defense to different application scenarios.

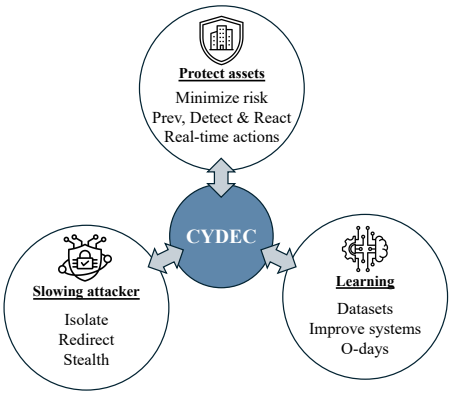

Figure 1. Cyber Deception uses case and main advantages.

The idea behind CYDEC is to create a digital environment based on deception that makes it difficult for attackers to

ISBN:978-84-09-62140-8

accomplish their malicious goals and makes them believe that they have gained access to valuable data or systems. This proactive approach not only seeks to defend against attacks by slowing them down or even neutralizing the attack within a controlled environment but also to gather information about the tactics and motivations of adversaries that allows organizations to strengthen their defenses based on the knowledge gained. CYDEC is becoming an essential tool in the modern cybersecurity arsenal, helping businesses and government entities improve their resilience in the face of ever-evolving cyber threats. As an example, the use of honeypots could be used for the detection and study of threats, in order to improve detection systems. [12].

However, analyzing vast volumes of heterogeneous data from various sources regarding attacker/threat behaviors poses a significant challenge. While humans can perform this task manually, it's laborious and time-consuming, prompting the emergence of Artificial Intelligence (AI). AI enhances prevention, detection, and reaction systems, enabling continuous learning and improvement in security mechanisms. Machine learning algorithms enable AI to process extensive data and identify patterns indicative of malicious activity. Furthermore, AI's adaptive nature enables security systems to promptly respond to emerging threats.

The convergence between CYDEC and Artificial Intelligence (AI) promises to mark a revolution in cybersecurity in the future [14]. As the AI continues to develop and mature, its integration with the CYDEC will become increasingly promising and crucial in the fight against cyber threats. In the near future, we can anticipate that AI will drive the creation of even more sophisticated and adaptable CYDEC environments. In fact, AI systems can generate and deploy decoys and traps autonomously and in real-time, making cyber attacks much more difficult to execute successfully. The AI will also enable continuous, in-depth analysis of cybercriminals' tactics and strategies, which will help organizations anticipate and defend against emerging threats such as "0-days" (vulnerabilities that have just been discovered and do not yet have a patch to address them) more effectively. In addition, the AI and CYDEC will work simultaneously to improve the detection and mitigation of cyber threats. In particular, AI based systems will be able to identify patterns of suspicious behavior in real-time and make decisions.

Despite the great advantage of the union of these two aspects, there are no papers that identify the union of these two concepts and identify the gaps and opportunities of the interconnection of the two.

In the light of the above, the most promising innovations in the paper are summarized below:

- We give an insight into the most important technical principles and their associated tools to help future researchers choose and implement methods of CYDEC.
- We contextualize the current use of CYDEC and the main advantages and disadvantages of its use by organizations. We also show which are its main application scenarios.
- We identified gaps and opportunities in the union of CYDEC and AI to focus researchers on key areas of exploration, improving CYDEC techniques and leveraging the benefits of AI to improve security systems.

The reminder of this article is as follows: a compilation of CYDEC's main techniques as well as its main tools are presented in Section II. In Section III, we provide a contextualization of CYDEC and its main application scenarios. In the Section IV, we detail the main gaps and opportunities that CYDEC and AI together present and, finally, in Section V, we review the main conclusions of the work and propose possible next steps.

II. Related work

Next, we are going to define which are the papers related to CYDEC. These works will be divided according to the technique used, and subsequently, the main tools used to carry out each technique will be identified and defined.

A. Techniques

This section will explore various CYDEC techniques used to deceive, disorient, and deter adversaries. Each technique offers an additional layer of protection to strengthen an organization's security, from the implementation of honeypots and honeynets to the application of obfuscation and perturbation.

- **Honeypots** [15]: They are computer systems designed to simulate real resources and attract attackers. Honeypots are configured to record and analyze attacker behaviors, which help organizations to understand the tactics and techniques used by adversaries and detect suspicious activity within the organizations' network.
- **Honeynets** [15]: They are networks of interconnected honeypots. While an individual honeypot can simulate a single system, a honeynet is a collection of honeypots that simulate an entire network.
- **Honeytokens** [15]: They are lures of information placed within a system or network to lure attackers and alert them to suspicious activity. Honeytokens can be files, fake credentials, or any other type of false data that attackers can find and indicate an intrusion when used or accessed.
- **Honeyfiles** [15]: These are fictitious files or documents that contain information valuable to attackers. By accessing these files, attackers reveal their presence and inform defenders about their targets and tactics.
- **Redirections** [16]: It is a technique that involves redirecting attackers' network traffic to fake or isolated systems, away from real critical systems and data. This diverts attackers' efforts and allows defenders to detect and respond to threats.
- **Decoys** [17]: These are fake systems, applications, or data designed to resemble an organization's real assets. The main difference with honeypots is that a decoy should look as much like a real system as possible.
- **Obfuscation** [18]: The process of hiding or masking systems or data's true purpose or operation. In cybersecurity, obfuscation makes it difficult for attackers to understand and analyze assets.
- **Perturbation** [19]: It involves introducing deliberate changes to an organization's systems, networks, or data to confuse attackers and hinder their activities. Disruption can include changes to network topology, system configuration, or data distribution.

- **Moving Target Defense (MTD)** [20]: It is a strategy that involves continually changing an organization's security parameters to make it more difficult for attackers to predict and exploit vulnerabilities.

B. Tools

Understanding key CYDEC techniques should be paired with identifying the essential tools for executing those techniques effectively. Below, we identify the main tools with which to carry out CYDEC and, in addition, we can see in Table I the relationship of each tool with the techniques that can be implemented.

Table I
LIST OF TECHNIQUES AND CORRESPONDENT TOOLS IDENTIFIED IN THE CYDEC CONTEXT.

Techniques	Tools
Honeypots	Awesome Honeypots
Honeynets	Modern Honey Network (MHN) Specter
Honeytokens	Canarytokens Metta Honeybits
Honeyfiles	Canarytokens
Redirections	HoneyPorts Iptables Fiddler
Decoys	Clonezilla Acronis True Image
Obfuscation	Obfsproxy
Perturbation	Packet generator tool
MTD	UBER MIRAGE DESIR

- **Awesome Honeypots** [21]: A curated list of awesome honeypots, plus related components and much more, divided into categories such as Web, services, and others, with a focus on free and open source projects.
- **MHN** [22]: It is an open-source platform for centralized honeypot management and threat data collection. MHN facilitates the configuration and deployment of honeypots and provides tools for analyzing and sharing threat data.
- **Specter** [23]: It is a commercial solution that provides a distributed honeypot network and an intuitive user interface for analyzing threat data. Specter helps organizations detect and respond to cyber threats by creating digital decoys.
- **Canarytokens** [24]: It is a tool that allows generating honeytokens, such as web links, documents, and email addresses, to detect unauthorized activity. Canarytokens alerts administrators when one of these decoys is accessed or used, indicating a possible intrusion.
- **Metta** [25]: It is an open-source tool that generates honeytokens and helps track the spread of insider threats. Metta can be used to simulate the presence of sensitive data and detect unauthorized access attempts.
- **Honeybits** [26]: It is a tool that generates false data or decoys within a network to deceive attackers and discourage attacks. Honeybits can be used to simulate the presence of valuable assets and lure attackers away from real network resources.
- **Honeyports** [27]: They are fake ports or spoofed services that lure attackers and log their activities. Honeyports can be implemented using firewall software or network tools to fool adversaries and protect real systems.
- **Iptables** [28]: It is a packet filtering and Network Address Translation (NAT) tool on Linux-based operating systems. iptables can be used to redirect network traffic and control access to network services and resources.
- **Fiddler** [29]: It is an HTTP traffic debugging tool that allows network administrators to analyze and manipulate web communications. It can be used to create deception scenarios and detect malicious activity on the network.
- **Clonezilla** [30]: It is an open-source disk cloning tool that allows creating and restoring disk images of complete systems. Clonezilla can be used to create identical systems.
- **Acronis True Image** [31]: It is a commercial backup and recovery solution that offers advanced system cloning features. Acronis True Image can be used to clone systems in real-time.
- **Obfsproxy** [32]: It is a tool used to obfuscate network traffic and complicate the work of attackers. Obfsproxy can be combined with anonymizing software like Tor to hide behavior and activities.
- **Packet generator tool** [33]: It is a tool that allows creating and sending customized data packets through a network to add changes in the network behavior to deceive the attacker and not to obtain behavioral patterns.
- **UBER, MIRAGE, DESIR** [34]: These three tools are part of the same research group focused on BAT. Depending on the scenario where we want to perform, we can choose one of them.

As discussed in this Section and Table I, we have identified tools with functionality associated with a technique of CYDEC to help future researchers find the easiest way to implement different CYDEC strategies. Any techniques described above can be implemented using programming languages, software-defined networking (SDN), or similar. In addition, the table identifies the most commonly used tools to carry out CYDEC techniques, leaving aside tools that are rarely used, paid tools or those that do not offer the required level of support and updating.

Existing works in the literature focus on defining and identifying CYDEC techniques, leaving aside how they could be implemented or executed. Due to this lack of knowledge, the compilation and association undertaken in this work represent a novelty that will furnish future researchers with the necessary knowledge to execute these CYDEC techniques.

III. WHAT IS CYBER DECEPTION?

CYDEC, as a cyber defense strategy, represents a fascinating and underexplored facet of the digital security world. In a landscape where cyber threats are constantly evolving, deception is an intriguing and promising way to protect critical assets. While its implementation may be limited, the

convergence of the CYDEC with the new capabilities of the AI offers an exciting prospect for its future development and application. In Section III-A, we discuss a review on the current use of CYDEC, while in Section III-B we show the main CYDEC application scenarios.

A. Actual use of CYDEC and AI

Although the concept of CYDEC has been under development for several years, its practical application, especially when combined with AI, remains relatively limited. This is due to several factors, including the technical complexity of implementing effective CYDEC strategies and the need to overcome concerns about potential unwanted side effects, such as false positives or interference with legitimate operations [35].

Despite these challenges, there is a growing recognition that CYDEC, combined with AI, can provide a highly effective layer of defense against various threats [36]. The AI provides advanced analytics capabilities that can detect patterns and anomalies in data at a speed and scale that are impossible for humans.

When used with CYDEC, AI (CYDEC-AI) can further enhance organizations' ability to prevent, detect, and mitigate security threats. For example, deception systems can improve their methods with the continuous learning that takes place thanks to the AI. By performing this action, we will be transforming our system into an autonomous system with the ability to operate independently.

In the current era, there are several companies with extensive experience in using deception-based technologies [37], some of which are based in Spain and have direct ties to U.S. government funding agencies. These types of companies seek to use the CYDEC to mitigate threats they detect in their systems more sophisticatedly. The main disadvantage of the vision of these companies is that they focus all their effort on the CYDEC, leaving aside the advantage of using the AI to improve these systems autonomously and to generate intelligence for each attack that is made towards our organization. An additional benefit of integrating artificial intelligence into these systems is the ability to adapt and learn from past attacks, enabling a faster and more effective response to new threats. In addition, AI can automate detection and response processes, freeing up human resources and reducing threat detection and mitigation time.

Another important aspect of researchers' current use of CYDEC is that all these years in which CYDEC has been present, much emphasis has been placed on the Honey-X part. At the same time, the other techniques with great potential to prevent, detect, and mitigate threats and attacks have been forgotten. As a result, we have obtained a great knowledge of the Honey-X part but a great lack of knowledge of redirection or decoy strategies with an important potential.

Figure 2 shows a high-level architecture detailing the linking components between CYDEC and the AI. In addition, the application scenarios identified in Section III-B are shown. In the Figure we see how CYDEC-AI is able to perform defensive actions in both prevention, detection and reaction to threats by complementing the various advantages of CYDEC and the use of AI to enhance the decision and performance techniques.

B. Application scenarios

Another key point when studying the concept of CYDEC is where defense strategies based on CYDEC can be carried out, i.e., in which scenarios it can be applied effectively. Below, we show the main application scenarios in which CYDEC techniques can be used to improve defense:

- **Critical Infrastructure:** In sectors such as energy, transportation, healthcare, and utilities, CYDEC-AI can protect critical infrastructure against targeted cyber attacks. Implementing CYDEC systems powered by CYDEC-AI can detect and mitigate emerging threats [38], ensuring the availability and integrity of vital services to society. For this purpose, networks of decoy traps could be created to trick the attacker into not detecting which of the systems is the real one.
- **Defense and Homeland Security:** CYDEC-AI plays a crucial role in protecting military and government networks against advanced cyber threats. The use of various techniques in this area can greatly improve military defensive capabilities as well as those of public institutions by improving and innovating their security techniques. In this context, the use of decoy-driven redirection techniques would be of great use along with improved information gathering and attribution techniques [39].
- **Internet of Things (IoT) environments:** In the IoT context, CYDEC-AI can protect IoT devices and networks against intrusions and cyber attacks. In order to deceive attackers in such scenarios, MTD strategies can be used to make the attack more difficult while not consuming unnecessary resources [40].
  By implementing CYDEC techniques and behavioral analytics supported by CYDEC-AI, organizations can detect and mitigate threats in real-time, ensuring the security and integrity of IoT ecosystems.
- **5G/6G:** In next-generation networks, such as 5G and future 6G, CYDEC-AI plays a key role in protecting the communications infrastructure from threats. By utilizing CYDEC techniques and vulnerability analysis driven by AI, organizations can mitigate the security risks associated with these emerging technologies, ensuring fast and secure connectivity for users and devices. In this way, any attacker found in wireless networks can be tricked or moved to a different network so as not to harm the connectivity of other users [41].

As we progress in the digital age, the adoption and evolution of CYDEC in partnership with AI are crucial to keep us one step ahead of cyber adversaries. In this direction, collaboration between industry, academia, and government agencies will be critical to unlocking the full potential of these powerful tools and safeguarding the integrity and security of our digital infrastructures.

Moreover, acknowledging that CYDEC is predominantly utilized in network and cybersecurity realms, it's pivotal to consider its potential extension to diverse contexts. This includes detecting financial fraud, safeguarding online privacy, countering social media disinformation, and curtailing public opinion manipulation campaigns. Expanding into new domains demands interdisciplinary teamwork and innovative adaptation of existing techniques, promising substantial

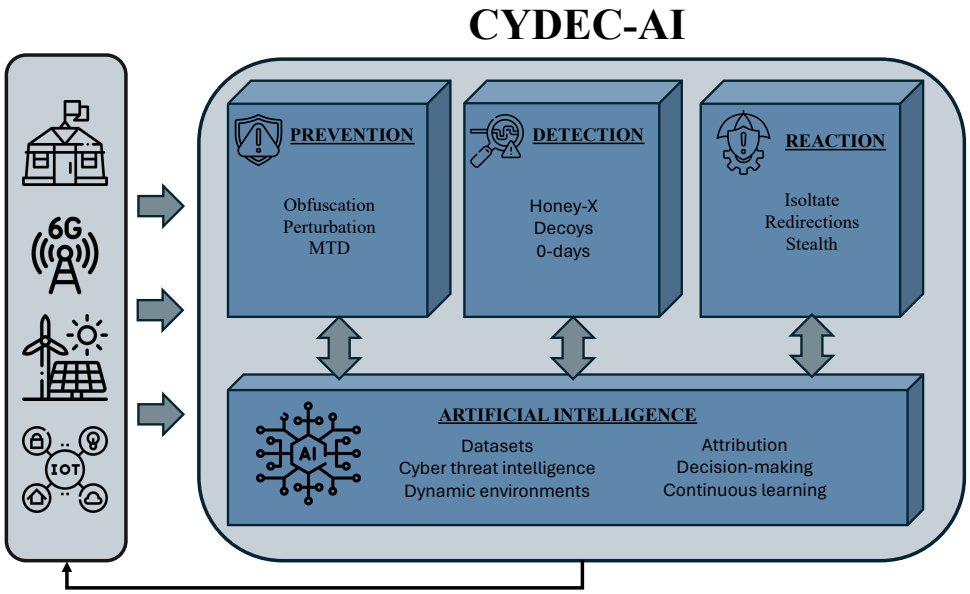

Figure 2. High level architecture CYDEC-AI highlighting the main defensive actions it can perform.

progress in digital threat defense.

C. Ethics aspects

Ethics in the field of CYDEC [42] emerges as a crucial area for reflection and action in the digital age. In this context, the ethical challenges are diverse and require detailed consideration to ensure that the use of these techniques does not violate fundamental values or generate unintended consequences.

One of the main ethical issues facing the field of CYDEC is how to ensure that activities conducted through systems such as CYDEC adhere to sound moral and legal standards. It is crucial to clearly define under what circumstances it is ethically justifiable to use disinformation or data manipulation, and how we can ensure that these practices do not violate individual or collective rights, such as privacy or freedom of expression [43].

Ensure thorough examination of these crucial inquiries to steer the conscientious advancement and application of these technologies. Additionally, prioritize transparency and accountability in CYDEC techniques' utilization. Users should openly disclose their actions and goals, and be governed by robust oversight and control measures.

This implies not only responsibility on the part of individual actors [44], but also the need for clear policies and regulations governing the use of CYDEC in different contexts, whether military, corporate or governmental.

Furthermore, incorporating real-world instances illustrating ethical dilemmas in CYDEC enhances comprehension of the field's ethical challenges. It's crucial to consider various viewpoints, encompassing effects on both individuals and society, to comprehensively tackle these issues. By recognizing these challenges, we can advance towards a more ethical and accountable utilization of CYDEC techniques, fostering a safer digital environment for everyone.

IV. CYDEC-AI: What is still missing?

As previously mentioned, the intersection between CYDEC and AI has emerged as a crucial field in cybersecurity, offering new perspectives and approaches to address evolving cyber threats. However, as we progress with this technological convergence, many challenges and opportunities emerge that must be addressed to realize the full potential of CYDEC-AI.

In this section, we explore a variety of key gaps and opportunities in CYDEC-AI. From the need for more accurate metrics to adapt to increasingly sophisticated attacks, each challenge presents an opportunity to innovate and improve our cyber defenses. In addition, we examine how the utilization of AI can provide innovative solutions to address these problems, from continuous learning to multiple attack detection.

Throughout this exploration, we will identify key areas where additional research and collaboration between industry, academia, and government agencies is required to advance the field of CYDEC-AI and ensure cyberspace security in an increasingly digitized and connected world.

A. Gaps

First of all, we will define the most important generic gaps. We will also define what each gap consists of and the associated opportunities within these gaps. We can see a summary of the relationship between gaps and opportunities in Table II.

1) Improving CYDEC techniques: Continuous enhancement of CYDEC techniques is vital to combat the evolving cyber threat landscape. Many existing techniques rely on established principles and may not adequately address sophisticated cyber attacks. A key challenge is their limited adaptability to emerging adversary tactics, such as advanced evasion techniques and exploitation of specific system vulnerabilities. Moreover, the lack of diversity in deception techniques can render decoys and deception environments easily identifiable by attackers, diminishing CYDEC effectiveness. This gap in

Table II
LIST OF GAPS-OPPORTUNITIES IDENTIFIED IN THE CONTEXT OF CYDEC-AI.

Gaps	Opportunities
Improvement of the techniques of CYDEC	Designing a decision making approach to generate environments of CYDEC Implement CYDEC techniques with stealth behavior. Update and improve attack modeling.
Integration of mechanisms for AI to CYDEC	Leveraging AI for continuous learning. Improve attack attribution techniques. Researching and adopting other, newer AI techniques. Develop and provide appropriate datasets.
Joining of different CYDEC methods	Propose a comprehensive framework integrating prevention, detection, and response strategies with CYDEC
Evaluation of implemented solutions	Identify additional specific metrics. Develop and employ more advanced test environments. Implement realistic testbeds.

improving CYDEC techniques exposes organizations to more complex and harder-to-detect cyber threats, jeopardizing data and system integrity, confidentiality, and availability.

2) Integration of AI mechanisms into the CYDEC: The effective integration of AI mechanisms into CYDEC solutions is essential to strengthen the cyber defenses of organizations. However, despite the enormous potential offered by AI, its full implementation and adoption in the CYDEC domain still faces significant challenges. One of the main limitations is the lack of adequate infrastructure and resources to develop, deploy, and maintain complex AI systems in cybersecurity environments. In addition, AI requires extensive and high-quality datasets to train models accurately, which can be difficult to obtain in dynamic and constantly changing cyber environments. This also includes the significant breakthroughs in recent advancements in AI models and mechanisms, where the field has evolved greatly, and AI strategies must be adapted to these new AI methods. The lack of integration of AI mechanisms can limit the ability of CYDEC solutions to proactively adapt to emerging threats and detect and mitigate cyber attacks more accurately and effectively. The main gap in the use of AI in CYDEC environments is given by the constant improvement of reaction and detection systems as well as the improvement of decision making. In addition, there is also a big gap in the use of AI techniques for attack attribution and the creation of more sophisticated deception environments.

3) Joining different methods of CYDEC: The use of different methods of CYDEC for the improvement of defense capabilities presents as a determining gap. However, there is currently a gap in the integration and coordination of these methods. The lack of unification can result in redundancies, gaps in defense coverage, and difficulties in managing and maintaining multiple CYDEC solutions. In addition, the lack of a unified approach can make it difficult to effectively detect and respond to cyber attacks, as adversaries can exploit gaps between different defense methods. This gap in bridging different methods of security leads to a weakness in acting effectively against attacks/threats.

4) Evaluation of implemented solutions: The correct evaluation of CYDEC solutions is still a task to be solved in order to identify which of the techniques implemented in different scenarios is the most determinant. However, there is currently a gap in the ability of organizations to conduct comprehensive and regular evaluations of these solutions. The lack of a systematic and standardized approach can result in incomplete or superficial assessments, making it difficult to identify vulnerabilities and weaknesses in cyber defenses. In addition, the lack of clear and objective metrics in environments where CYDEC and AI are used is a weakness in the implementation of both.

B. Opportunities

Once the most important gaps have been defined, we will define and develop the associated opportunities, relating each opportunity with the gap to which it contributes, as reported in Table II.

1) Designing a decision making approach to generate environments of CYDEC: Developing a structured and systematic approach to decision-making in generating CYDEC environments could provide an opportunity to improve the effectiveness of these solutions. By establishing clear criteria and well-defined decision-making processes, the likelihood of creating more effective and difficult-to-detect decoys and deception environments could be increased.

2) Implement CYDEC techniques with stealth behavior: Developing CYDEC techniques with stealthy behavior could provide an opportunity to improve the ability to defend against cyber attacks. By operating unobtrusively and going undetected by adversaries, these techniques could increase the effectiveness of CYDEC solutions by deceiving attackers and protecting systems undetected. Stealth is the most important feature of any CYDEC technique, and currently, few papers focus on obtaining that feature.

3) Update and improve attack modeling: Updating and improving attack modeling to accurately reflect modern tactics and techniques used by cyber adversaries could provide an opportunity to strengthen cyber defenses. By developing more advanced and up-to-date models, the ability of CYDEC solutions to anticipate and respond to emerging threats can be improved, increasing the effectiveness of protection against cyber attacks. In this context, the use of CYDEC for the creation of information gathering environments plays a key role in accomplishing this task.

4) Leveraging AI for continuous learning: Taking full advantage of the AI for continuous learning could provide an opportunity to improve the capability of CYDEC solutions.

By enabling solutions to learn and adapt to new threats continuously, the ability to detect and respond to evolving cyber attacks could be improved. In other words, making the best use of the environments created in response to feed our detection and prevention systems. Using different AI techniques in real-time to improve our systems will equip our organizations with a continuous defense, i.e., every new threat that appears we will be able to detect and mitigate it. For example, by isolating new threats in controlled environments we will be able to observe their behaviors and observe patterns that will help us to detect them.

5) Improve attack attribution techniques: Improved attack attribution techniques in CYDEC environments, supported by AI, could provide an opportunity to strengthen cyberattack response capabilities. The development of more accurate and effective methods for attributing attacks could make it easier to identify the culprits and take appropriate measures to mitigate damage. AI can play a crucial role in this process by analyzing large volumes of security event data, identifying patterns and correlations between different attacks and helping to determine the origin and *modus operandi* of attackers.

6) Researching and adopting other, newer AI techniques: Researching and adopting other newer AI techniques, such as Natural Language Processing (NLP), deep learning, evolutionary computing, federated learning, generative AI, etc., could open up new possibilities in the CYDEC field. These techniques could improve the ability of solutions to detect and respond to cyber threats more effectively while providing greater flexibility and adaptability to ever-changing cybersecurity environments. Thanks to techniques such as federated learning, we will be able to share the intelligence generated by different organizations to improve common systems. In addition, the emergence of generative AI will make the deception scenarios created increasingly sophisticated.

7) Develop and provide appropriate datasets: Developing and providing appropriate datasets to feed AI systems in the CYDEC field could provide an opportunity to improve the ability to detect and respond to cyber attacks. By having relevant and representative datasets, the accuracy and effectiveness of CYDEC solutions in identifying and mitigating threats could be improved. This type of action can be developed thanks to technologies such as honeypots, honeynets, or decoys. A large deployment of machines in the cloud capable of collecting information and tagging it automatically can be carried out.

8) Propose a comprehensive framework integrating prevention, detection, and response strategies with CYDEC: Developing a comprehensive framework that integrates prevention, detection, and response strategies with CYDEC could provide an opportunity to improve the effectiveness of cyber defenses. Such a framework would enable more effective coordination between different cybersecurity components and optimize the use of CYDEC to protect systems against various threats.

9) Identify additional specific metrics: Developing and establishing additional specific metrics to evaluate the effectiveness of CYDEC solutions could provide an opportunity to improve the understanding and measurement of their impact on cybersecurity. These metrics could enable a more accurate assessment of solution performance and facilitate the identification of areas for improvement. In particular, it will be necessary to obtain Key Performance Indicators (KPIs) of the quality of the different CYDEC mechanisms for each technique, i.e., a study of the most decisive characteristics of CYDEC techniques should be carried out to identify their key points and to obtain the most appropriate metrics.

10) Develop and employ more advanced test environments: Developing and employing more advanced test environments that simulate realistic cyberattack scenarios could provide an opportunity to improve the CYDEC solutions. These environments would allow researchers and developers to test and refine solutions against more complex threats, resulting in increased robustness and adaptability of CYDEC's tools in detecting and mitigating advanced cyber attacks. The main idea is to replicate past attack scenarios to observe the behavior of CYDEC techniques in critical situations. We will also have to launch Advance Persistence Threats (APTs) against this type of techniques to observe the ability of CYDEC techniques to mitigate this type of attacks.

11) Implement realistic testbeds: Developing and establishing realistic testbeds could provide a solid platform for testing and continuously improving CYDEC solutions. This would enable more effective validation of existing techniques and open opportunities for developing new strategies and tools in a controlled environment, which could lead to significant advances in cyber defense. Unlike the previous opportunity, this one is based on the creation and testing of CYDEC techniques in more realistic environments with more concurrent data and more users and systems.

As we have seen in this section, there are a number of gaps and opportunities related to the use of CYDEC and AI in the cybersecurity domain. While these opportunities are accompanied by significant challenges, they also offer a wide range of benefits that have the potential to significantly improve the effectiveness of our defense mechanisms.

V. Conclusions

This paper proposed a contextualization of the current state of CYDEC, exploring its main techniques and tools, application scenarios, and more. It also joining the concepts of AI with CYDEC, aiming to revolutionize security benefits. To achieve this, we identify the primary gaps in using CYDEC and its linkage with AI and define the key opportunities to fill them.

This research shed light on the current role of CYDEC in the security arsenal of organizations and enterprises, and the lack of a common understanding of its key concepts. It also underscores the urgent need to harness the power of AI in the context of CYDEC. These insights reveal numerous opportunities to unlock the vast potential that remains untapped. This research provides a robust foundation to drive the development and implementation of innovative strategies that fully leverage CYDEC and security capabilities.

The path forward involves a comprehensive compilation of the state of the art on CYDEC, unifying key concepts of its many features and creating a taxonomy. We also envision the development of a generic framework that can unite all CYDEC techniques for use in threat prevention, detection, and mitigation. Lastly, we advocate further exploring the opportunities identified through their application to leverage their potential and generate comprehensive defense tools.

Acknowledgements

This work has been partially funded by the strategic projects CDL-TALENTUM and DEFENDER from the Spanish National Institute of Cybersecurity (INCIBE) by the Recovery, Transformation, and Resilience Plan, Next Generation EU.

This work has also received funding from the European Defence Fund (EDF) under grant agreement No 101103044 EDF-2021-CYBER-R-CDAI-2.

Disclaimer

Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union (EU). The EU cannot be held responsible for them.

References

[1] The Internet: evolution and growth statistics. Stackscale B.V. (2024, February 6). https://www.stackscale.com/blog/internet-evolution-statistics/

[2] Sagiroglu, S., & Sinanc, D. (2013, May). "Big data: A review". In *2013 international conference on collaboration technologies and systems* (CTS) (pp. 42-47). IEEE. https://doi.org/10.1109/CTS.2013.6567202

[3] Singh, N. K., & Mahajan, V. (2021). "Analysis and evaluation of cyber-attack impact on critical power system infrastructure". *Smart Science*, 9(1), 1-13. https://doi.org/10.1080/23080477.2020.1861502

[4] Kaur, J., & Ramkumar, K. R. (2022). "The recent trends in cyber security: A review". *Journal of King Saud University-Computer and Information Sciences*, 34(8), 5766-5781. https://doi.org/10.1016/j.jksuci.2021.01.018

[5] Huang, K., Siegel, M., & Madnick, S. (2018). "Systematically understanding the cyber attack business: A survey". *ACM Computing Surveys* (CSUR), 51(4), 1-36. https://doi.org/10.1145/3199674

[6] Del Campo, E. A. P., Polo Alvis, S., Sánchez Acevedo, M. E., & León Quiroga, A. (2023). "Cyberspace: a new frontier. In Frontiers–Law", *Theory and Cases* (pp. 89-126). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-13607-8_5

[7] García Cid, M. I., Gil Pérez, M., Jorquera Valero, J. M., López Martínez, A., Maestre Vidal, J., Martínez Pérez, G., ... & Sotelo Monge, M. A. (2023). "European framework and proofs-of-concept for the intelliGent aUtomAtion of cybeR Defence Incident mAnagemeNt". En *2023 JNIC Cybersecurity Conference* (JNIC), pp. 483-484. IEEE. https://doi.org/10.23919/JNIC58574.2023.10205559

[8] P. Nespoli, D. Papamartzivanos, F. G. Mármol, & G. Kambourakis. (2017). "Optimal Countermeasures Selection Against Cyber Attacks: A Comprehensive Survey on Reaction Frameworks." In *IEEE Communications Surveys & Tutorials*, 20(2), 1361-1396. https://doi.org/10.1109/COMST.2017.2781126

[9] Stabili, D., Romagnoli, R., Marchetti, M., Sinopoli, B., & Colajanni, M. (2022, April). "Exploring the consequences of cyber attacks on powertrain cyber physical systems". In *2022 International Conference on Control, Robotics and Informatics (ICCRI)* (pp. 96-103). IEEE. https://doi.org/10.1109/ICCRI55461.2022.00023

[10] Almeshekah, Mohammed H. and Spafford, Eugene H. (2016), "Cyber Security Deception", *Springer International Publishing*, 23-50, https://doi.org/10.1007/978-3-319-32699-3_2

[11] Medenou Choumanof, R. D. (2022). "Cyberspace Defense Operations from a MILDEC Perspective" (Master's thesis).

[12] Matin, I. M. M., & Rahardjo, B. (2020, October). "The use of honeypot in machine learning based on malware detection: A review". In *2020 8th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-6). IEEE. https://doi.org/10.1109/CITSM50537.2020.9268794

[13] Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). "Artificial intelligence for cybersecurity: Literature review and future research directions". *Information Fusion*, 101804. https://doi.org/10.1016/j.inffus.2023.101804

[14] Quanyan Zhu. (2023). "The Doctrine of Cyber Effect: An Ethics Framework for Defensive Cyber Deception". *arXiv*. 2302.13362. https://doi.org/10.48550/arXiv.2302.13362

[15] Qin, X., Jiang, F., Cen, M., & Doss, R. (2023). "Hybrid cyber defense strategies using Honey-X: A survey". *Computer Networks*, 109776. https://doi.org/10.1016/j.comnet.2023.109776

[16] Lopez, P. B., Nespoli, P., & Gil Pérez, M. (2024). "Cyber Deception Reactive: TCP Stealth Redirection to On-Demand Honeypots". *arXiv preprint* arXiv:2402.09191. https://doi.org/10.48550/arXiv.2402.09191

[17] Ferguson-Walter, K. J., Major, M. M., Johnson, C. K., & Muhleman, D. H. (2021). "Examining the efficacy of decoy-based and psychological cyber deception". In *30th USENIX security symposium (USENIX Security 21)* (pp. 1127-1144).

[18] Urias, V. E., Stout, W. M., Luc-Watson, J., Grim, C., Liebrock, L., & Merza, M. (2017, October). "Technologies to enable cyber deception". In *2017 International Carnahan Conference on Security Technology (ICCST)* (pp. 1-6). IEEE. https://doi.org/10.1109/CCST.2017.8167793

[19] Pawlick, J., Colbert, E., & Zhu, Q. (2019). "A game-theoretic taxonomy and survey of defensive deception for cybersecurity and privacy". *ACM Computing Surveys (CSUR)*, 52(4), 1-28. https://doi.org/10.1145/3337772

[20] Gao, C., Wang, Y., Xiong, X., & Zhao, W. (2021, June). "Mtdcd: an mtd enhanced cyber deception defense system". In *2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)* (Vol. 4, pp. 1412-1417). IEEE. https://doi.org/10.1109/IMCEC51613.2021.9482133

[21] Awesome Honeypots. paralax. https://github.com/paralax/awesome-honeypots

[22] Modern Honey Network. Pwnlandia/MHN: Modern Honey Network. GitHub. https://github.com/pwnlandia/mhn

[23] Specter. Specter intrusion detection system. http://gabiam.com/software/laura_chapelle/Software/specter/

[24] Canarytokens. Thinkst Canary. https://help.canary.tools/hc/en-gb/articles/4701687447325-What-are-Canarytokens

[25] Uber-Common. Metta. GitHub. https://github.com/uber-common/metta

[26] 0x4D31. Honeybits. GitHub. https://github.com/0x4D31/honeybits

[27] Securitygeneration. Honeyport. GitHub. https://github.com/securitygeneration/Honeyport

[28] Iptables. https://www.acens.com/comunicacion/wp-content/images/2014/07/wp-acens-iptables.pdf

[29] Fiddler. Telerik.com. https://www.telerik.com/fiddler

[30] Team., D. Clonezilla. https://clonezilla.org/

[31] Acronis true image. Acronis. https://www.acronis.com/es-es/support/trueimage/2021/

[32] The Tor Project, Inc. Tor. Tor Project: Pluggable Transports. https://2019.www.torproject.org/docs/pluggable-transports

[33] Kt. Packet generator tool. Packet Generator Tool. https://www.netscantools.com/nstpro_packet_generator.html

[34] UBER, MIRAGE, DESIR. Sun Security lab. Available at: https://sunlab-gmu.github.io/research/mtd.html (Accessed: 08 March 2024).

[35] Zhu, M., Anwar, A. H., Wan, Z., Cho, J. H., Kamhoua, C. A., & Singh, M. P. (2021). "A survey of defensive deception: Approaches using game theory and machine learning". *IEEE Communications Surveys & Tutorials*, 23(4), 2460-2493. https://doi.org/10.1109/COMST.2021.3102874

[36] Lopes Antunes, D., & Llopis Sanchez, S. (2023, August). "The Age of fighting machines: the use of cyber deception for Adversarial Artificial Intelligence in Cyber Defence". In *Proceedings of the 18th International Conference on Availability, Reliability and Security* (pp. 1-6). https://doi.org/10.1145/3600160.3605077

[37] Cyber deception systems. Acalvio. (2023, July 25). https://www.acalvio.com/resources/analyst-reports/cyber-deception-systems-cds-market-spotlight/

[38] Morozov, D. S., Vakaliuk, T. A., Yefimenko, A. A., Nikitchuk, T. M., & Kolomiiets, R. O. (2023, April). "Honeypot and cyber deception as a tool for detecting cyber attacks on critical infrastructure". In *Proc. 3rd Edge Comput. Workshop Doors*.

[39] Steingartner, W., & Galinec, D. (2021). "Cyber threats and cyber deception in hybrid warfare". *Acta Polytechnica Hungarica*, 18(3), 25-45.

[40] Ge, M., Cho, J. H., Kim, D., Dixit, G., & Chen, I. R. (2021). "Proactive defense for internet-of-things: moving target defense with cyberdeception". *ACM Transactions on Internet Technology (TOIT)*, 22(1), 1-31. https://doi.org/10.1145/3467021

[41] Adebayo, A., & Rawat, D. B. (2020, January). "Deceptor-in-the-middle (ditm): Cyber deception for security in wireless network virtualization". In *2020 IEEE 17th annual consumer communications & networking conference (CCNC)* (pp. 1-6). IEEE. https://doi.org/10.1109/CCNC46108.2020.9045164

[42] Ethics. National Cyber Deception Laboratory. https://www.cyberdeception.org.uk/ethics/

[43] Zhu, Q. (2023). "The Doctrine of Cyber Effect: An Ethics Framework for Defensive Cyber Deception". *arXiv preprint* arXiv:2302.13362.

[44] Yadav, R. (2021). "Social deception in online platform: concept, attacks and ethical issues". *INFORMATION TECHNOLOGY IN INDUSTRY*, 9(3), 321-326.