International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Use of Artificial Intelligence on Cyber Security and the New-Generation Cyber-attacks

Sumit Kumar Das[1], Payal Panda[2]

[1]MS in Cybersecurity, Massachusetts Institute of Technology, USA
[2]MS in Artificial Intelligence, IU International University, Germany

Abstract:
Cybersecurity is a fast growing and evolving discipline that is always in the news over the last decade, as the number of threats rises and cybercriminals constantly endeavour to stay a step ahead of law enforcement. Over the years, although the original motives for carrying out cyberattacks largely remain unchanged, cybercriminals have become increasingly sophisticated with their techniques.
Not only have there been a lot more cyberattacks in recent years, but they have also gotten much more advanced. Therefore, developing a cyber-resilient strategy is most significance. In the event of a cyberattack, traditional security measures are insufficient to prevent data leaks. Cybercriminals have mastered the use of cutting-edge methods and powerful tools for data intrusion, hacking, and assault.
In this we are proposing applications of artificial intelligence (AI) technology in the creation of intelligent models for securing systems against attackers. AI technologies can quickly advance to meet complicated problems, making them useful as fundamental cybersecurity tools to identify malware attacks, AI-based systems can provide efficient and robust cyber security against phishing and spam emails, network intrusions, and data breaches capabilities and alert the security during the impact. Here, we explore AI's potential in improving cybersecurity solutions, by identifying both its strengths and weaknesses. We also discuss future research opportunities associated with the development of AI techniques in the cybersecurity field across a range of application domains.

Keywords: Cybersecurity, Artificial Intelligence, Machine Learning

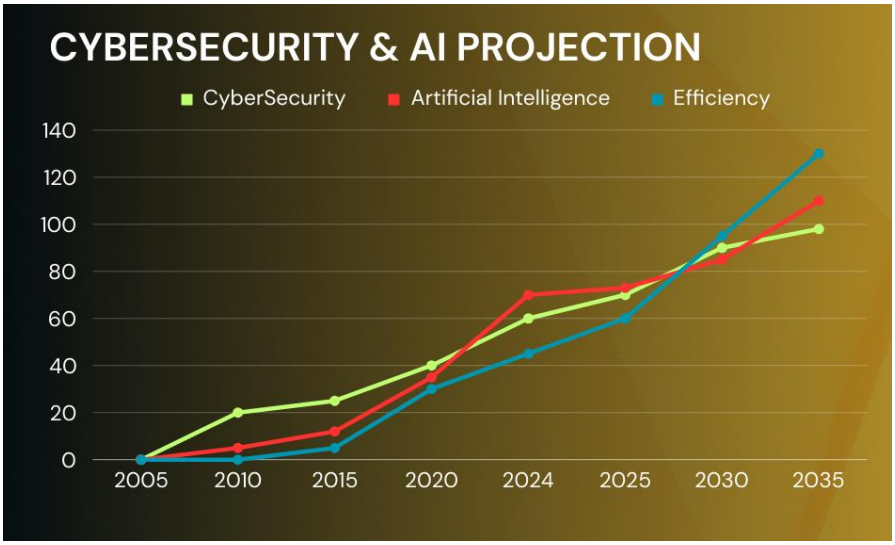

Fig 1: year wise demand growth for AI and Cybersecurity

SECTION I.

Introduction

Cybersecurity is the practice of protecting systems, networks, and programs from digital attacks. These cyberattacks are usually aimed at accessing, changing, or destroying sensitive information; extorting money from users via ransomware; or interrupting normal business processes. Cybersecurity is defined as a set of processes, human behaviour, and systems that help safeguard electronic resources.

According to Moore's law that forecasts the doubling of components on an integrated circuit every two years (along with decreasing costs associated with chip manufacturing), cybercriminals are increasingly doubling the effectiveness of their attack tools for half the cost every few months. Global cybersecurity spending is expected to exceed $1 trillion from 2017 to 2021, where spending on cybersecurity already increased by almost 40 percent from 2013 to $66 billion.

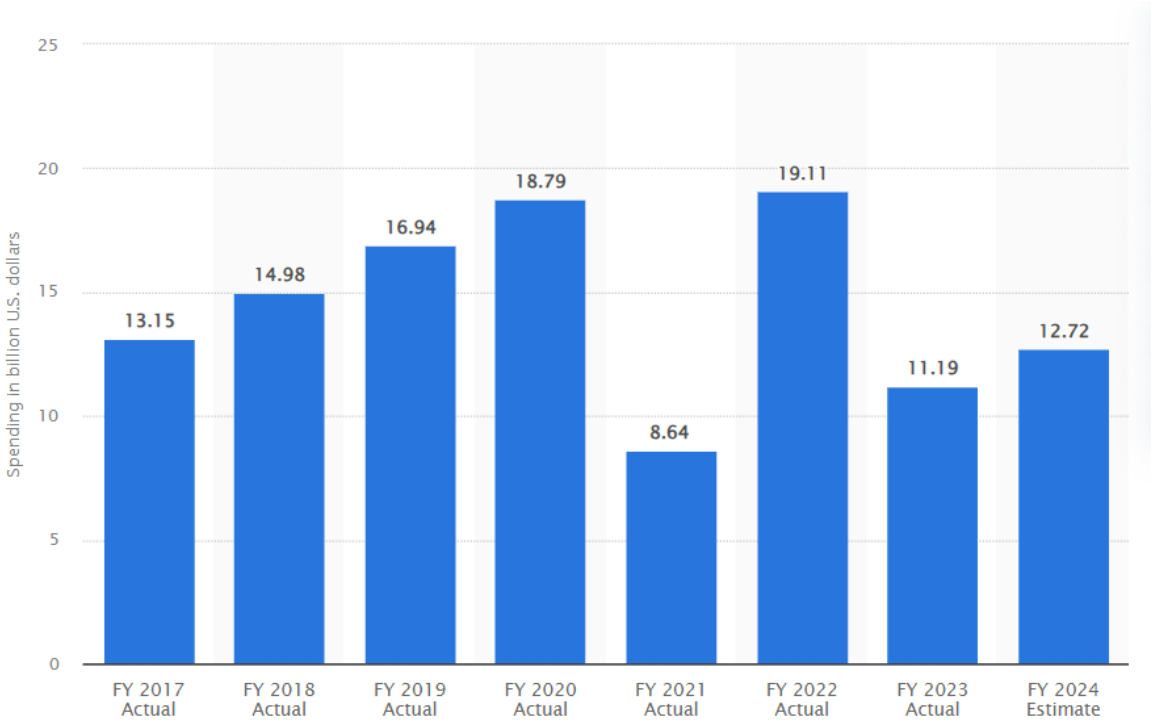

Fig 2: ***Estimated budget of the U.S. government for cybersecurity in FY 2017 to 2024(in billion U.S. dollars)***

In the past few years, cybersecurity researchers have started to explore Artificial Intelligence (AI) approaches to improve cybersecurity. Likewise, cybercriminals are also using AI to launch increasingly sophisticated, and complex cyberattacks while hiding their tracks. However, in this work, we focus on how AI-based cybersecurity solutions could protect us from attackers better and minimize or prevent data breaches.

Advances in AI have led to many exciting research results and systems since its emergence in the 1950s. Further developments led to the emergence of machine learning and deep learning. Today, AI has been deployed in far-reaching application areas, including healthcare, agriculture, space, law, and manufacturing. The continuous performance improvements in computer hardware and software (along with their decreasing costs), coupled with new paradigms such as big data and cloud computing, have led to the development and deployment of a wide range of AI systems with varying capabilities.

Today, many of these AI systems now perform a broad range of complex tasks that include learning, planning, problem solving, decision making, and face/speech recognition. Since the 1980s, another major development within the AI field has been the emergence of machine learning technologies that help computer systems learn and adapt to various conditions by using their past experiences, patterns, and knowledge. Recently, we have witnessed an increasing interest in the use of AI and machine learning techniques to fight cyberattacks. A strong motivation for the use of these techniques stems from the large amounts of data that are constantly being produced today, which requires significant resources and time to analyse and detect any patterns, anomalies, or intrusions in traffic data.

In a recent report by Juniper research, the authors predict that the cost of cybersecurity incidents will increase from 3 trillion each year to more than 5 trillion in 2024, an average yearly growth of 11 percent.

The key sources of cyberthreats include:

- **Script kiddies:** These are novices who have trained to create cyberattack tools to hack into vulnerable computing systems and to make a quick buck or boost their ego through such activities.
- **Criminal organizations:** These include those involved in illegal operations, who launch cyberattacks that can cause a Denial of Service (DoS), steal data or state secrets because of data breaches, seek payments through ransomware, and so on.
- **Nation states:** This involves state-sponsored cybercriminal activities perpetrated against enemy nations with the intent of crippling the victim nation's economy or critical infrastructures, causing fatalities, disruption of state-sponsored programs, or to ultimately topple the government.
- **Terrorists:** They attempt to cause nationwide losses and major disruptions to society's critical infrastructures, such as causing massive power outages in a victim country through cyberattacks.
- **Spies (including business rivals):** They steal trade secrets to gain an unfair market advantage.
- **Disgruntled employees:** Employees who are stressed and unhappy with their jobs, rifts with management, or other factors may attempt to cause financial or reputation losses to the organization by carrying out a cyberattack against corporate resources.
- **External attackers and insider threats:** Experts with a strong knowledge about the operation of computing resources as well as human behaviour, who attempt to exploit vulnerable systems and gain (mainly financially) through such acts or simply cause major disruptions to the organization's normal operations.

One type of threat that's becoming more prevalent and continuously evolving in complexity over the years is the zero-day threat which has not been previously seen by cybersecurity or software/hardware development staff. Consequently, the attacker exploits the computing resources' security vulnerability (software or hardware) the same day it becomes known. When a zero-day attack targets a software vulnerability, the patching of the security hole must be initiated from the software developer or vendor as quickly as possible. Such security patches take time to be created and rolled out on a global scale. During this interim period, all non-patched systems are exposed to the cyberthreat of the zero-day vulnerability.

An example of such a threat is zero-day malware that can easily penetrate a target system while bypassing malware detection software such as anti-virus. Cybercriminals are using advanced techniques for code obfuscation, defined as concealment of malicious code within "legitimate-appearing code" that

can be delivered to a victims' system in the form of an email attachment. Naïve users may open these attachments or click an embedded link to a malicious website, leading to system compromise and more severe consequences—including data held for ransom, compromise, and even sensitive data disclosure. Hidden malware within ads that appear on legitimate websites are also a clever technique for compromising end-user systems through zero-day exploits. Even the most up-to-date security software will not be able to detect obfuscated code embedded within such adware.

The German AV-TEST GmbH research institute for IT security registers more than 350,000 new malware programs and potentially unwanted applications every day. In fact, in 2019, the institute identified more than 140 million new malware programs, which translates to an equivalent of 266 types of malwares every minute.

As the complexity of cyberthreats increases, the key drivers pushing for increased cybersecurity at the corporate level include:

- **Lack of cyber governance skills at the C-level**: Executives such as the Chief Information Security Officer (CISO) and the Chief Information Officer (CIO), do not easily make changes in security strategy at the corporate level. Such changes would safeguard corporate resources against the ever evolving and dynamic nature of cyber threats of contemporary times. The aggravating factor is the fact that cyber criminals are not privy to C-Level culture of organizations, and therefore cybersecurity is increasingly posing a concern at executive meetings.
- **Opportunities to harness state-of-the-art cybersecurity detection techniques:** Current computing systems become more efficient in data crunching, while at the same time the data required for cybersecurity analysis has become available. This trend has advanced cybersecurity analysis techniques such as machine learning, data mining, and knowledge discovery. Data mining is a subcomponent of knowledge discovery, where a specific sequence of steps is applied to data with the intent of extracting patterns. In addition, knowledge discovery also comprises data cleaning, selection, and the application of prior knowledge and established techniques for interpreting the results extracted. Machine learning and data mining significantly overlap, as they employ similar methods and processes. Whilst machine learning focuses on classification of data samples and prediction of events or behaviours, data mining focuses on the discovery of previously unseen patterns in data (very much like detection of zero-day cyberattacks). The advancement of these techniques has become one of the key drivers for organizations to achieve their goals, including their cybersecurity vigilance.
- **Fragmented cybersecurity frameworks:** Despite having a plethora of frameworks for securing an organization's resources against cyberthreats, the choice remains a largely difficult question for an organization's cybersecurity decision makers. Some industries such as the insurance sector do not have a proper reference model to follow to ensure the requisite cybersecurity. This is attributed mainly to the lack of consumer data to build legitimate and illicit profiles, upon which machine learning or AI techniques can be applied; definitions of fraud differ between the insurance sector and the banking sector. In the former case, insurers mainly worry about policies being opened without a priori customer knowledge, and they operate in a fragmented regulatory environment. For instance, unlike banking, the insurance industry is not tightly regulated in the US, consequently encumbering the adoption of silver-bullet cyber prevention strategies because they invariably depend upon regulation. Therefore, the industry-specific cybersecurity framework, or lack thereof, hinders the

realization of cybersecurity goals in a wide range of industries. A similar concern arises in Supervised Control and Data Acquisition (SCADA) systems that comprise a range of commercial off-the-shelf hardware and software and rely upon standardized communication protocols. While integrity and availability are important cybersecurity concerns for SCADA systems, confidentiality is secondary. Precedence is typically given to safety, reliability, robustness, and maintainability of such systems, and therefore security takes a backseat.

Research contributions of this work

We summarize the main contributions of this work as follows:

- We present an overview of the cybersecurity threat landscape and discuss traditional security solutions (i.e., non-AI based solutions) that have been used to protect from the various threats.
- We discuss the weaknesses of traditional cybersecurity solutions and describe how emerging AI solutions can improve cybersecurity.
- Finally, we present some key challenges faced by the cybersecurity community that must be addressed in the future.

SECTION II.

Cybersecurity Threats and Legacy Cybersecurity Solutions

Over the last decade, many types of cyberthreats have emerged. Next, we briefly review those threats. According to a recent report, the top 10 cyberthreats we face today include:

1. Malware

Malware — or malicious software — is any program or code that is created with the intent to do harm to a computer, network, or server. Malware is the most common type of cyberattack, mostly because this term encompasses many subsets such as ransomware, trojans, spyware, viruses, worms, keyloggers, bots, cryptojacking, and any other type of malware attack that leverages software in a malicious way.

- **Ransomware:** In a ransomware attack, an adversary encrypts a victim's data and offers to provide a decryption key in exchange for a payment. Ransomware attacks are usually launched through malicious links delivered via phishing emails, but unpatched vulnerabilities and policy misconfigurations are used as well.
- **Fileless Malware:** Fileless malware is a type of malicious activity that uses native, legitimate tools built into a system to execute a cyber-attack. Unlike traditional malware, fileless malware does not require an attacker to install any code on a target's system, making it hard to detect.
- **Spyware:** Spyware is a type of unwanted, malicious software that infects a computer or other device and collects information about a user's web activity without their knowledge or consent.
- **Adware:** Adware is a type of spyware that watches a user's online activity to determine which ads to show them. While adware is not inherently malicious, it has an impact on the performance of a user's device and degrades the user experience.
- **Trojan:** A trojan is malware that appears to be legitimate software disguised as native operating system programs or harmless files like free downloads. Trojans are installed through social engineering techniques such as phishing or bait websites. The Zeus trojan malware, a variant, has the goal accessing financial information and adding machines to a botnet.
- **Worms:** A worm is a self-contained program that replicates itself and spreads its copies to other computers. A worm may infect its target through a software vulnerability, or it may be delivered via

phishing or smishing. Embedded worms can modify and delete files, inject more malicious software, or replicate in place until the targeted system runs out of resources.

- **Rootkits:** Rootkit malware is a collection of software designed to give malicious actors control of a computer network or application. Once activated, the malicious program sets up a backdoor exploit and may deliver additional malware. Rootkits take this a step further by infecting the master boot prior to the operating system being on boot up, going undetectable at times.
- **Mobile Malware:** Mobile malware is any type of malware designed to target mobile devices. Mobile malware is delivered through malicious downloads, operating system vulnerabilities, phishing, smishing, and the use of unsecured Wi-Fi.
- **Exploits:** An exploit is a piece of software or data that opportunistically uses a defect in an operating system or an app to provide access to unauthorized actors. The exploit may be used to install more malware or steal data.
- **Scareware:** Scareware tricks users into believing their computer is infected with a virus. Typically, a user will see scareware as a pop-up warning them that their system is infected. This scare tactic aims to persuade people into installing fake antivirus software to remove the "virus." Once this fake antivirus software is downloaded, then malware may infect your computer.
- **Keylogger:** Keyloggers are tools that record what a person types on a device. While there are legitimate and legal uses for keyloggers, many uses are malicious. In a keylogger attack, the keylogger software records every keystroke on the victim's device and sends it to the attacker.
- **Botnet:** Botnet is a network of computers infected with malware that are controlled by a bot herder. The bot herder is the person who operates the botnet infrastructure and uses the compromised computers to launch attacks designed to crash a target's network, inject malware, harvest credentials, or execute CPU-intensive tasks.
- **MALSPAM:** Malicious malware (MALSPAM) delivers malware as the malicious payload via emails containing malicious content, such as virus or malware infected attachments.

2. Denial-of-Service (DoS) Attacks

A Denial-of-Service (DoS) attack is a malicious, targeted attack that floods a network with false requests to disrupt business operations.

In a DoS attack, users are unable to perform routine and necessary tasks, such as accessing email, websites, online accounts, or other resources that are operated by a compromised computer or network. While most DoS attacks do not result in lost data and are typically resolved without paying a ransom, they cost the organization time, money, and other resources to restore critical business operations.

The difference between DoS and Distributed Denial of Service (DDoS) attacks has to do with the origin of the attack. DoS attacks originate from just one system while DDoS attacks are launched from multiple systems. DDoS attacks are faster and harder to block than DOS attacks because multiple systems must be identified and neutralized to halt the attack.

3. Phishing

Phishing is a type of cyberattack that uses email, SMS, phone, social media, and social engineering techniques to entice a victim to share sensitive information — such as passwords or account numbers — or to download a malicious file that will install viruses on their computer or phone.

- **Spear Phishing:** Spear-phishing is a type of phishing attack that targets specific individuals or organizations typically through malicious emails. The goal of spear phishing is to steal sensitive information such as login credentials or infect the targets' device with malware.
- **Whaling:** A whaling attack is a type of social engineering attack specifically targeting senior or C-level executive employees with the purpose of stealing money or information, or gaining access to the person's computer to execute further cyberattacks.
- **SMiShing:** Smishing is the act of sending fraudulent text messages designed to trick individuals into sharing sensitive data such as passwords, usernames, and credit card numbers. A smishing attack may involve cybercriminals pretending to be your bank or a shipping service you use.
- **Vishing:** Vishing, a voice phishing attack, is the fraudulent use of phone calls and voice messages pretending to be from a reputable organization to convince individuals to reveal private information such as bank details and passwords.

4. Spoofing

Spoofing is a technique through which a cybercriminal disguises themselves as a known or trusted source. In so doing, the adversary can engage with the target and access their systems or devices with the goal of stealing information, extorting money, or installing malware or other harmful software on the device.

Spoofing can take different forms, which include:

- **Domain Spoofing:** Domain spoofing is a form of phishing where an attacker impersonates a known business or person with fake website or email domain to fool people into the trusting them. Typically, the domain appears to be legitimate at first glance, but a closer look will reveal subtle differences.
- **Email Spoofing:** Email spoofing is a type of cyberattack those targets businesses by using emails with forged sender addresses. Because the recipient trusts the alleged sender, they are more likely to open the email and interact with its contents, such as a malicious link or attachment.
- **ARP Spoofing:** Address Resolution Protocol (ARP) spoofing or ARP poisoning is a form of spoofing attack that hackers use to intercept data. A hacker commits an ARP spoofing attack by tricking one device into sending messages to the hacker instead of the intended recipient. This way, the hacker gains access to your device's communications, including sensitive data.

5. Identity-Based Attacks

CrowdStrike's findings show that 80% of all breaches use compromised identities and can take up to 250 days to identify.

Identity-driven attacks are extremely hard to detect. When a valid user's credentials have been compromised and an adversary is masquerading as that user, it is often very difficult to differentiate between the user's typical behaviour and that of the hacker using traditional security measures and tools.

Some on the most common identity-based attacks include:

- **Kerberoasting:** Kerberoasting is a post-exploitation attack technique that attempts to crack the password of a service account within the Active Directory (AD) where an adversary masquerading as an account user with a service principal name (SPN) requests a ticket, which contains an encrypted password, or Kerberos.

- **Man-in-the-Middle (MITM) Attack:** A man-in-the-middle attack is a type of cyberattack in which an attacker eavesdrops on a conversation between two targets with the goal of collecting personal data, passwords, or banking details, and/or to convince the victim to take an action such as changing login credentials, completing a transaction or initiating a transfer of funds.
- **Pass-the-Hash Attack:** Pass the hash (PtH) is a type of attack in which an adversary steals a "hashed" user credential and uses it to create a new user session on the same network. It does not require the attacker to know or crack the password to gain access to the system. Rather, it uses a stored version of the password to initiate a new session.
- **Golden Ticket Attack:** In a golden ticket attack, adversaries attempt to gain unlimited access to an organization's domain by accessing user data stored in Microsoft Active Directory (AD) by exploiting vulnerabilities in the Kerberos identity authentication protocol. This allows adversaries to bypass authentication methods.
- **Silver Ticket Attack:** A silver ticket is a forged authentication ticket often created when an attacker steals an account password. A forged service ticket is encrypted and enables access to resources for the specific service targeted by the silver ticket attack.
- **Credential Harvesting:** In credential harvesting, cybercriminals gather user credentials — such as user IDs, email addresses, passwords, and other login information to then access systems, gather sensitive data, or sell it in the dark web.
- **Credential Stuffing:** Credential stuffing attacks work on the premise that people often use the same user ID and password across multiple accounts. Therefore, possessing the credentials for one account may be able to grant access to other, unrelated account.
- **Password Spraying:** The basics of a password spraying attack involve a threat actor using a single common password against multiple accounts on the same application. This avoids the account lockouts that typically occur when an attacker uses a brute force attack on a single account by trying many passwords.
- **Brute Force Attacks:** A brute force attack is using a trial-and-error approach to systematically guess login info, credentials, and encryption keys. The attacker submits combinations of usernames and passwords until they finally guess correctly.
- **Downgrade Attacks:** Downgrade attacks are a cyberattack where adversaries take advantage of a system's backward compatibility to force it into less secure modes of operation, such as forcing a user to go into a HTTP version of a website instead of HTTPS.

6. Code Injection Attacks

Code injection attacks consist of an attacker injecting malicious code into a vulnerable computer or network to change its course of action. There are multiple types of code injection attacks:

- **SQL Injection:** A SQL Injection attack leverages system vulnerabilities to inject malicious SQL statements into a data-driven application, which then allows the hacker to extract information from a database. Hackers use SQL Injection techniques to alter, steal or erase application's database data.
- **Cross-Site Scripting (XSS):** Cross Site Scripting (XSS) is a code injection attack in which an adversary inserts malicious code within a legitimate website. The code then launches as an infected script in the user's web browser, enabling the attacker to steal sensitive information or impersonate

the user. Web forums, message boards, blogs and other websites that allow users to post their own content are the most susceptible to XSS attacks.

- **Malvertising:** Malvertising attacks leverage many other techniques, such as SEO poisoning, to carry out the attack. Typically, the attacker begins by breaching a third-party server, which allows the cybercriminal to inject malicious code within a display ad or some element thereof, such as banner ad copy, creative imagery, or video content. Once clicked by a website visitor, the corrupted code within the ad will install malware or adware on the user's computer.

7. Supply Chain Attacks

A supply chain attack is a type of cyberattack that targets a trusted third-party vendor who offers services or software vital to the supply chain. Software supply chain attacks inject malicious code into an application in order to infect all users of an app, while hardware supply chain attacks compromise physical components for the same purpose. Software supply chains are particularly vulnerable because modern software is not written from scratch: rather, it involves many off-the-shelf components, such as third-party APIs, open-source code, and proprietary code from software vendors.

8. Insider Threats

IT teams that solely focus on finding adversaries external to the organization only get half the picture. Insider threats are internal actors such as current or former employees that pose danger to an organization because they have direct access to the company network, sensitive data, and intellectual property (IP), as well as knowledge of business processes, company policies or other information that would help carry out such an attack.

Internal actors that pose a threat to an organization tend to be malicious in nature. Some motivators include financial gains in exchange for selling confidential information on the dark web, and/or emotional coercion using social engineering tactics, such as pretexting or business email compromise (BEC) attacks. On the other hand, some insider threat actors are not malicious in nature but instead are negligent in nature. To combat this, organizations should implement a comprehensive cybersecurity training program that teaches stakeholders to be aware of any potential attacks, including those potentially performed by an insider.

9. DNS Tunnelling

DNS Tunnelling is a type of cyberattack that leverages domain name system (DNS) queries and responses to bypass traditional security measures and transmit data and code within the network.

Once infected, the hacker can freely engage in command-and-control activities. This tunnel gives the hacker a route to unleash malware and/or to extract data, IP, or other sensitive information by encoding it bit by bit in a series of DNS responses.

DNS tunnelling attacks have increased in recent years, in part because they are relatively simple to deploy. Tunnelling toolkits and guides are even readily accessible online through mainstream sites like YouTube.

10. IoT-Based Attacks

An IoT attack is any cyberattack that targets an Internet of Things (IoT) device or network. Once compromised, the hacker can assume control of the device, steal data, or join a group of infected devices

to create a botnet to launch DoS or DDoS attacks.
[According to the Nokia Threat Intelligence Lab, connected devices are responsible for nearly one-third of mobile network infections – more than double the amount in 2019.]
Given that the number of connected devices is expected to grow rapidly over the next several years, cybersecurity experts expect IoT infections to grow as well. Further, the deployment of 5G networks, which will further fuel the use of connected devices, may also lead to an increase in attacks.

11. Eavesdropping attacks

These can be carried out by sniffing out the network communication line and misusing obtained data. Malicious actors may either passively sniff the line and obtain user data or actively attack the line, replacing messages with fictitious messages, and masquerade as legitimate users.

12. Birthday attacks

This hash of a message, also known as a message digest, which can be computed using a standard algorithm such as the Secure Hash Algorithm-1 (SHA-1). When this algorithm is applied to a message of arbitrary length, the output is a hash value of fixed length. The birthday attack refers to the attempt by a malicious actor to find two different messages that produce the same hash value. Consequently, the original message can be replaced with the other message that produces the same hash value, causing system and service disruption and data loss. Such attacks apply AI techniques to discover random messages that produce the same hash value as a legitimate message.

13. Network attacks

These are launched on the environment to disrupt services, steal individual/corporate data, and gain network intelligence. Malicious users exploit the Operating System's (OS's) weakness to gain access and tamper with the OS to achieve their malicious objectives. Some of these attacks are used to steal individual information, which can be used to gain access to individual/corporate data. In Table 1, we classified various network attacks based on their attack objectives, expected targeted device or application, data/ information exposed when specific attack is underway, type of environment affected when certain attacks occur, and how these attacks are detected.

Attack goal	Attack vector	Data exposure	Attack outcome	Environment	Attack detection
Stealing information	Hardware	Individual	Backdoor access; access to memory; Operating System (OS) tampering	Standalone device	Anomaly, signature
	Network	Centralized monitoring software; external 3rd party software	Corrupt device OS; exposure to Denial of Service (DoS) and Man in The Middle (MiTM) attack	Multiple devices	Anomaly
	Application, software	Email, Active Directory and application servers	Access to emails, personal Information, and various applications	Multiple devices and applications	Anomaly
	Media files	Individual	Access to personal data on computers and storage devices	Storage data	Anomaly
Tracking information	User credentials	Individual	Backdoor access; access to memory; Operating System (OS) tampering	Single & multiple users	Anomaly
	Application data	Individual	Protocols, IOS software control, DoS, DDoS and MiTM attacks	Application	Anomaly
	Monitoring user activities	Individual	Access to personal data	Single & multiple users	Anomaly
	Location data	Individual	Access to personal data	Single & multiple users	Anomaly
Device control	Hardware	Individual	Backdoor access; access to memory; Operating System (OS)	Single & multiple users	Anomaly, signature
	Network	Centralized monitoring software, external 3rd party software	Protocols, device control software, DoS, DDoS and MiTM attacks	Single & multiple devices	Anomaly
	Application, software	Centralized monitoring software, external 3rd party software	Protocols, general Input Output Software (IOS), software control, DoS, DDoS and MiTM attacks	Multiple devices and applications	Anomaly
	Location data	Individual	Access to personal data	Standalone device	Anomaly

Next, we briefly discuss traditional (non-AI) cybersecurity techniques for detecting cyberattacks:

1. *Game theory:* This has been previously applied to cybersecurity. The malicious actor is considered as one player in a game, and the victim's machine is the other player. Each player attempts to maximize his/her incentive through strategic movement, in which the player rationally justifies that the goal would be reached by the move. Each player's behaviours either can be known beforehand or remain concealed. An example of a game could be a smart grid environment where the attacker attempts to disrupt communication between a power system and a home, whereas the defender attempts to maintain connectivity between these various entities. At each step of the game, the attacker and the defender would adopt strategies to be successful in their respective goals.
2. *Rate control:* Attacks against the availability of systems include DoS and DDoS. Rate-control techniques can minimize the impact on such systems' operation when they are under attack by reducing the volume of incoming network traffic, through basic traffic throttling and redefining permission lists.
3. *Heuristics:* Firewalls and intrusion detection systems commonly rely on heuristics to identify the most apt rule for classifying network traffic as legitimate or anomalous. One such technique, performs a sequence of steps comprising substring matching to identify suspicious website addresses. The second phase of the presented scheme comprises the scanning of the web address through the OSINT websites, with the lowest score of the two scans considered for deciding on whether to let the data packets into the network or not.
4. *Signature-based intrusion detection:* A signature-based intrusion detection system makes use of a database that may store legitimate signatures corresponding to normal traffic or attack signatures corresponding to malicious traffic. The intrusion detection system matches the contents of incoming network packets with the stored signatures in real time. This technique's drawback is that in the absence of relevant signatures, intrusion detection systems are limited in their capabilities to accurately detect malicious traffic entering a network.
5. *Anomaly-based intrusion detection:* This technique creates a model of what can be perceived as the norm. The models can be in terms of rule-based policies, mathematical models, and statistical techniques. Deviations from the norm are regarded as attacks. When compared to the signature-based detection, such techniques have the advantage of being relieved from depending on signature patterns, thereby removing them from administrative efforts to collect signatures.
6. *Autonomous systems:* These have the capability to self-protect and self-heal, and to ensure reliability and availability, as in the case of the Bionic Autonomic Nervous System (BANS). This system is comprised of four different modules, namely, Cyber Neuron, Cyber Axon, Peripheral Nerve and Central Nerve. Cyber Neuron is used to protect against spyware and malware. Cyber Axon is an intelligent tool to recover from damage caused by spyware and malware. Similarly, Peripheral Nerve provides a robust defence against DoS/DDoS attacks by establishing a communication path between multiple cyber neurons deployed on different devices. Last, Central Nerve serves as a knowledge base against new attacks and to disseminate information to other security devices. Collaborative defence by peripheral nerves is proposed to block DoS and DDoS attacks through cooperation between devices within the network.
7. *End user security controls:* Current end-user devices such as mobile phones, smart portable devices (iPads), and personal computers require in-built security rather than add-ons. End users might not update their devices with the latest security patches, with some vendors attempting to push automatic

updates, to install security patches. The WannaCry ransomware attack is an example of an attack wherein the latest security patches provided by the vendor were not applied on all the end-user devices. Most of the time users are not aware of the implications of not applying the patches. In some cases, although some users may be aware of this fact, they do not either take the requisite action for securing their devices or they carry out incorrect procedures, exposing the devices through other vulnerabilities. A suggested control is to perform "out of sight" security, where automatic updates are pushed by vendors directly to end-user devices without the user's involvement. However, the challenge would be that software vendors must ensure that the security updates guard against new attacks (also known as zero-day attacks) and work seamlessly with all pre-existing software on the end-user device.

SECTION III.

Artificial Intelligence

AI is concerned with how machines can think or act correctly, given what they know. This universal definition includes how closely machines can think or act like humans (Fig. 3). At one end of the spectrum, machines are deemed to be intelligent if they can maximize the outcome on every state of the process. At the other end of the spectrum, the Turing Test sets the standard on machine intelligence. Under this test, a computer communicating with a human is said to have intelligence when the human cannot distinguish whether the responses come from a computer or a human. At both sides of the spectrum, AI embodies computing areas such as natural language processing, knowledge representation, logic, automated reasoning, machine learning, mathematics, and game theory. Early AI applications gave rise to thinking machines that solved puzzles such as geometry, checker games, and a family of blocks-world problems.

FIGURE 3. Spectrum on intelligent measures from thinking humanly through the Turing Test, to acting humanly to maximize the outcome.

After the proliferation of the Internet in the late '90s, software that behaved like humans gained popularity in terms of agent-based AI, commonly called bots. Ethical bots were made to spider the Internet for the benefit of search engines, yellow pages, and recommendation lists. They provide protection against unethical practices in Wikipedia articles where anybody can contribute as authors. In contrast, malicious bots also emerged to cheat in online games, post spams, and spread malware. In copying online games, bot programmers analysed the traffic flow between the game console and server to reverse engineer the game code. In posting spams, the bots mimicked the behaviour of human when online, such as surfing the pages before posting a message in a forum, rather than continuously posting messages. Malicious bots discourage cyber services to function properly, costing the service providers to have disheartened online visitors. As a result, some of the cybersecurity research investigated solutions that can detect and protect again malicious bots.

Studies found that game bots were active longer, were less social e.g. exchanging items or participating in an auction and have less variations in their sequence of actions when compared to human. Furthermore, game bots are more interested to collect items, while human players seek to collaborate with other players to complete challenges/quests. Similarly, spambots and malware bots can be detected from their behaviours being different than human, that can be detected through some distinctive communication patterns.

The most relevant AI applications to the cybersecurity area are in intrusion detection systems. Cybersecurity solutions often perform traffic analysis, where the Internet traffic is classified as either legitimate or malicious. At the dawn of the Internet, cyberattacks were identified with rule-based systems, where attacks could be detected based on their signatures. Over the years, as the number of Internet-connected devices and their applications increased, observing the huge amounts of network traffic being generated in real-time and creating rules which analyse this traffic have become time-consuming and make security protection systems behave defensively rather than proactively. Coupled with this advancing trend, technological advances are also benefiting attackers who are developing new sophisticated attack strategies that can avoid detection by current security systems. As the cyberthreat landscape continues to rise, we need advanced tools and technologies which can help detect, investigate, and make decisions faster for emerging threats. AI has the potential to intelligently analyse and automatically classify large amounts of Internet traffic. Today, cybersecurity solutions, based on ML technologies, are being used to automate the detection of attacks and to evolve and improve their capabilities over time. ML-based solutions are being used in intrusion detection systems as they can handle large volumes of data and a wide range of data attributes (e.g. many table columns) used for classification. Machine learning techniques learn from the collected Internet traffic to distinguish the malicious from the legitimate traffic class. It is worthwhile pointing out that due to the pervasiveness of machine learning in addressing cybersecurity issues, the adoption of the “machine learning” terminology has become interchangeable with “Artificial Intelligence” in the cybersecurity field.

A. Machine Learning

Conventionally, machine learning methods can be classified into two categories: supervised and unsupervised learning. In supervised learning, data samples are labelled according to their class (e.g., malicious, or legitimate). Training data, or data labelling is usually performed manually, requiring humans to detect data patterns with their classes. The trained data is input to an algorithm to create a mathematical model, which can output the predefined classes given new data samples. In unsupervised learning, no data labelling or training is required. Instead, the algorithms determine the degree of coherence/dispersion among data samples, systematically creating classes, and then classifying these samples according to the quality of data coherence within the class and data modularity between the classes.

However, discussions in machine learning blur the distinction between supervised/unsupervised machine learning algorithms. Mathematical, statistical, and probabilistic methods are used by machine learning techniques, allowing unsupervised algorithms to label the data used by supervised algorithms. This shows that taxonomy perspectives are converging, making it less essential to define machine learning algorithms based on whether they are supervised or unsupervised. Henceforth, we present an in-depth discussion of machine learning algorithms from a taxonomy perspective as described in, but in this

section, we discuss the predominant machine learning techniques that are effective for cybersecurity solutions.

Machine learning algorithms process data samples based on their determining factors, commonly called features. The data input is processed as a table of rows and columns, with rows serving as data samples and the columns representing their features. Naïve Bayes is a machine learning technique used to classify data based on the Bayesian theorem where the features are assumed to originate from independent events. The technique uses the computed probability of each class over all instances as the basis to find the probability of new data samples belonging to the class. Although the performance of Naïve Bayes classifiers degrades when more features come from dependent events, it is widely adopted, because it can inherently accept such a naïve assumption (that each feature comes from independent events) while still yielding acceptable results.

B. Decision Trees

A decision tree is a technique used to create a set of rules from the training data samples. The algorithm iteratively finds a feature that best categorizes data samples. The iterative division creates a sequence of rules for every side of the categories, resulting in a tree-like structure, until data samples with only one class are found after a division. Fig. 4 shows a decision tree example that classifies network traffic using rules that lead to normal or attack traffic classifications. The tree shows that, for example, if the flow of the traffic is low, but the duration of the traffic pattern is long, then it is classified as an attack. The technique provides an intuitive method for detecting cybersecurity issues, because it shows the result of a decision according to the feature values, as what is required by classifying observed events in cybersecurity as either legitimate or an attack. For example, the flow rate, size, and duration were used by decision trees to detect DoS attacks in addition to source/destination error rates. Furthermore, in detecting command injection attacks to robotic vehicles, decision trees were employed to categorize values from CPU consumption, network flow, and the amount of data written. This technique's benefit is that once the effective series of rules has been found, intrusion detection systems can classify Internet traffic in real time. The quality of generated real-time alerts is one of the most important attributes in detecting cyberattacks.

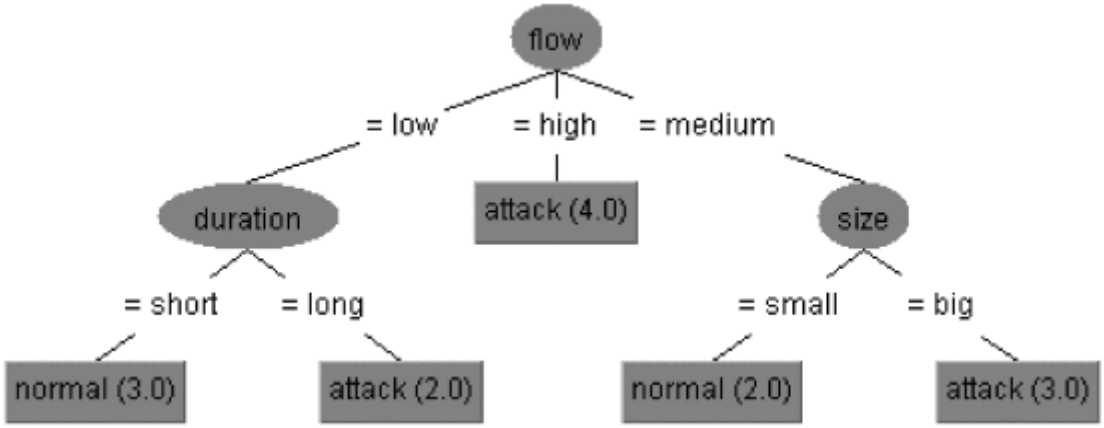

FIGURE 4 An example of a decision tree that classifies network traffic into attack and normal traffic type.

A different approach is the Rule-Learning technique, which seeks to find a set of feature values for each iteration while maximizing a score that defines the classification result's quality—for example, the

number of incorrectly classified data samples. Such an approach is like decision trees in that it generates a set of rules for classification. While decision trees find the best feature values that lead to a class, a rule-learning technique finds a set of rules that can describe a class. The advantage of a rule-learning technique is that it can factor human expert advice in generating rules. Consider a study that employed 28 features to detect DoS attacks in cloud networks. The features consisted of computer and network indicators, such as Input/Output (IO) reads, memory used, TCP flags detected, and the number of system resources opened. It generated a set consisting of rules derived from the features (e.g. IO_reads greater IO_reads(average)) and employed feature-ranking algorithms to discern the most relevant rules in finding the class. Afterward, the study employed human experts to optimize the rules, such as removing redundancies. Thus, the technique is suitable for intrusion detection systems where the configurations are mainly rule-based. Furthermore, the technique was generally employed as a performance benchmark to other machine learning techniques in detecting network intrusions.

C. K-Nearest Neighbours

The k-Nearest Neighbour (k-NN) technique learns from data samples to create classes or clusters. It was first proposed as a non-parametric pattern analysis to find the proportion of data samples in a neighbourhood that yields a consistent estimate of a probability. The neighbourhood was set as k-number of data samples according to a distance metric, usually the Euclidian distance to create clusters. The votes from all k neighbours decide how new data samples can be assigned to one of the clusters.

Fig. 5 illustrates the above technique. A new data sample (the red dot) was added to the data. In this example, the winning vote came from the highest number of data samples from one neighbouring cluster. Hence, when k = 3, the sample was put into Class 2. When k = 9, the sample was put into Class 1.

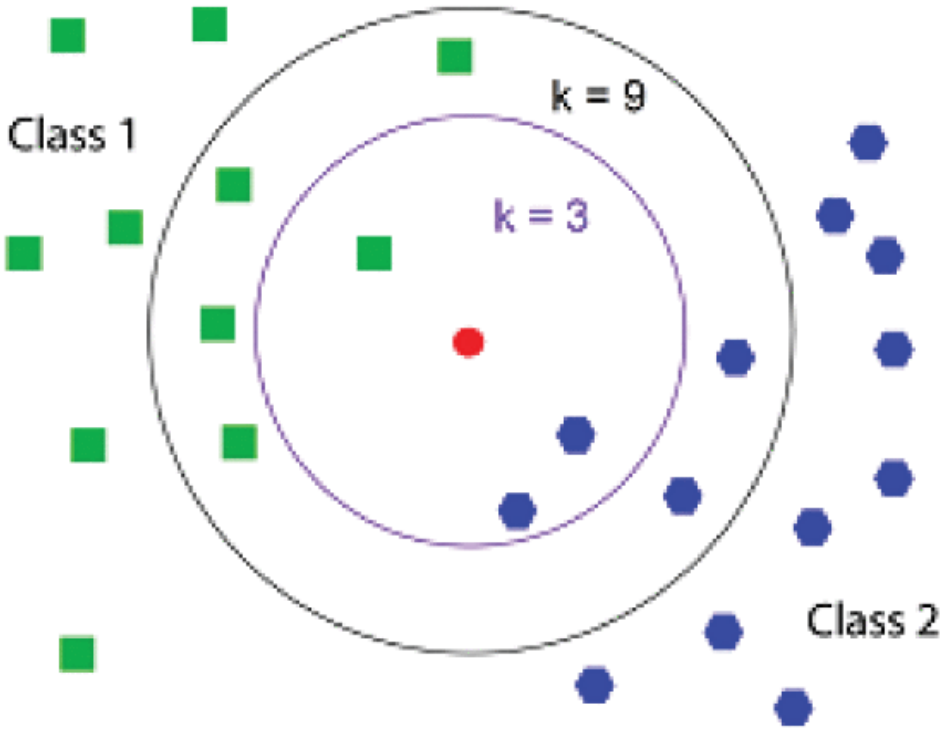

FIGURE 5. The k-Nearest Neighbour (k-NN) algorithm classifies data in class 1 and class 2, based on the k nearest data samples in the neighbourhood from the new data sample.

This technique is computationally complex even for small values of k. However, it is attractive for intrusion-detection systems because it can learn from new traffic patterns to reveal zero-day attacks as its unseen classes. Active research in this area thus seeks to find how k-NN can be used for real-time detections of cyberattacks. Recently, the technique was employed to detect attacks such as data tampering and false data injection against industrial control systems and smart grids. It performs well when the data can be represented through a model that allows the measurement of their distance to other data–for example, in terms of a Gaussian distribution or a vector.

D. Support Vector Machines

The Support Vector Machines (SVMs) technique extends linear regression models. While classifying data samples, SVMs find a plane that separates data samples into two classes (as shown in Fig. 6).

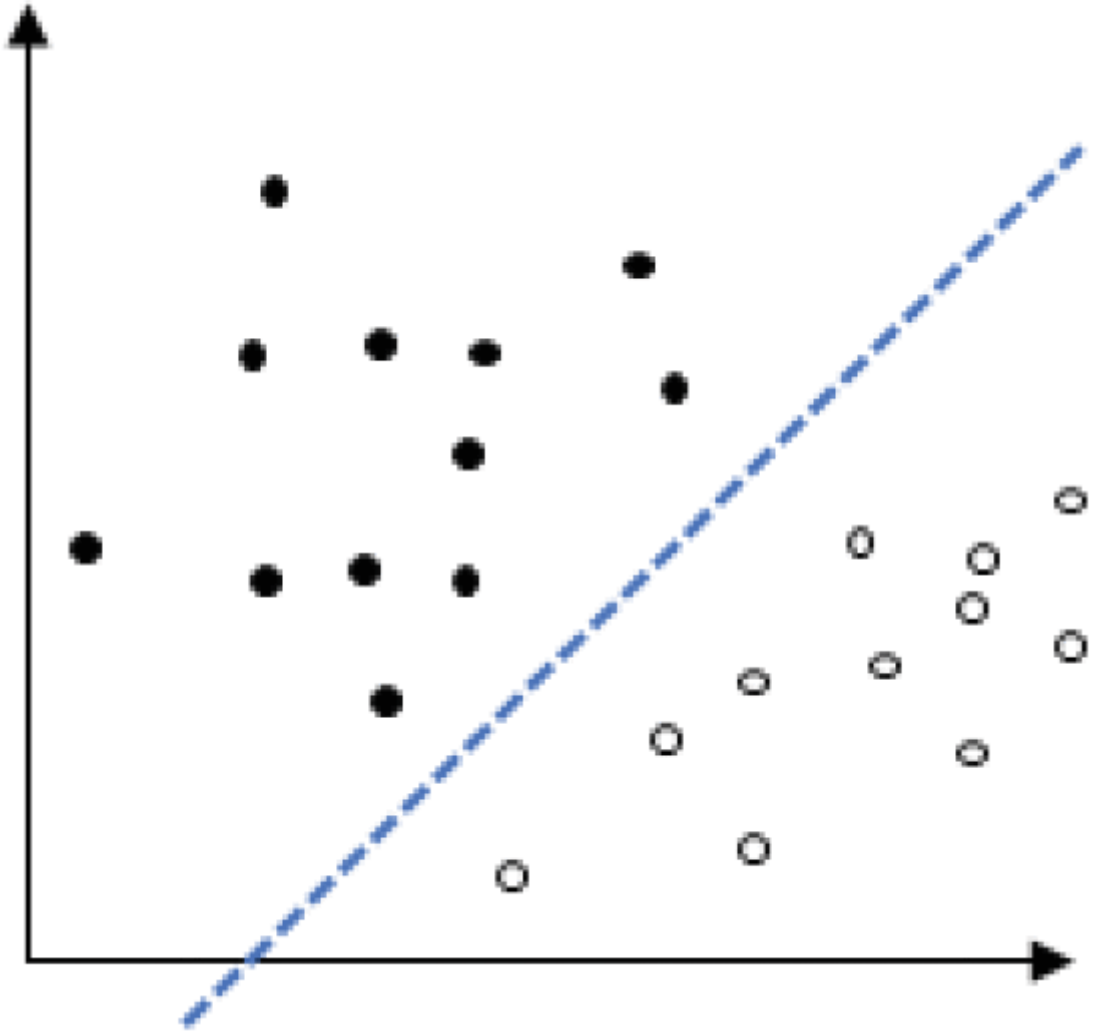

FIGURE 6. Support Vector Machines (SVMs) find a plane that separates data samples.

The separating plane can be shaped to form linear, non-linear, polynomial, Gaussian, Radial, sigmoid, and so on depending on the function employed (called a kernel). SVMs can also separate multiclass data (that is, not only data to be classified into two classes such as legitimate versus attack class as what the previous examples showed, but rather data to be classified into more than two classes) by employing more than one plane. This makes SVMs an attractive technique that can be used to analyse Internet

traffic patterns, which often consist of several classes such as HyperText Transfer Protocol (HTTP), File Transfer Protocol (FTP), Post Office Protocol 3 (POP3), and Simple Mail Transfer Protocol (SMTP).
SVM is a supervised machine learning technique, which requires training data to create a classification model. Therefore, it is used in applications where attacks can be simulated. For example, network traffic generated from the penetration testing conducted on a network system was used as the training data. SVM was employed to create a mathematical model to find a plane the penetration test traffic from normal traffic. A variation on its use creates a 1-class model for the normal traffic, while the model can be employed to detect anomalies when attack traffic was introduced. From these perspectives, the benefit of SMVs enables the development of attack detection models through simulations.

E. Artificial Neural Networks

The Artificial Neural Networks (ANNs) learning technique is inspired from how neurons in the brain work. ANN techniques model neurons in terms of a mathematical equation that reads a series of data samples to output a target value. The equation closely resembles the linear regression equation where data attributes of a sample are weighed to yield an output value. The ANN algorithm iterates until the output value is within the range of an acceptable error from the target value. In each iteration, the neurons learn by correcting their weights by measuring how far the error is from the target value, when given certain patterns identified from the data samples. When the error becomes negligible, the algorithm yields a mathematical equation that outputs an informative value such as the class, when given unseen data samples. ANN techniques can distinguish patterns that range from noisy to incomplete data samples. They are suitable for intrusion-detection systems because they adapt to new forms of communications.
In a cybersecurity study, an ANN application used the Cascade Correlation Neural Network (CCNN) which adds new hidden units to the hidden layer, step by step. When new events are detected, new hidden nodes are added to the network and only those are trained with the newly collected data thereby enabling a runtime adaptive and scalable system. In this work, the CCNN allows the training of the network with new data and does not need to retrain the whole network with the original data to learn from desktop-platform traffic patterns to detect port scanning to mobile networks. During the past decade, the rise of mobile devices has created new traffic patterns, causing previously built detection models obtained from desktop traffic to become obsolete. Port-scanning activities against mobile devices differed in their frequency of received packets and the number of ports scanned per second. The study showed that ANN port-scanning detection performance was comparable to other algorithms' performance, such as Decision Trees.
Another benefit of ANN is that it can detect zero-day attacks, because it can learn from recent incidents. For example, traffic patterns from having DoS attack incidents were fed to ANNs as the labelled training data, allowing the neurons to adjust their weights and detect unseen DoS attacks. When incidents such as DoS attacks occurred, the victim can testify that an attack has occurred, as opposed to other incidents (e.g., system penetration) where the attackers can cover their tracks, leaving the victim as gullible. Thus, ANNs is a suitable detection technique for cybersecurity applications where the attack class can be labelled when an incident (such as DoS) occurred, allowing the detection system to learn from the incident.

F. Self-Organizing Maps

Self-Organizing Maps (SOMs) take ANNs to the next level, namely, to self-adjust the neurons' weight to output a 2- or 3-dimensional (2D or 3D) map showing how the data can be grouped. The technique learns by finding the correlations that exist in data samples. Adjacent data samples share more similar features than the ones further away, thereby clustering data and providing an output in the form of a map. SOMs are computationally complex, making it unsuitable for real-time intrusion detection. Their major benefit lies in their ability to visualize the data, which is therefore useful in visualizing network anomalies. Without visualization, the outputs from intrusion-detection systems are hard to analyse. Visualization tools allow network operators to picture the normal pattern of traffic data (e.g., in terms of protocol interactions and traffic volume), thereby equipping them to effectively find anomalies in network traffic, including zero-day attacks. Although visualization approaches can point to anomalous events effectively, it still requires trained eyes to find anomalies in the data. Therefore, SOMs were employed as a complementary tool for detecting cyberattacks.

Since SOMs illustrates data in a 2D or 3D map, it is suitable to visualize multidimensional data (e.g., when the data in a table have many columns). In other words, SOMs reduce the dimensionality of data. Although there are other dimensional reduction techniques (such as Principal Component Analysis and Curvilinear Component Analysis), they do not visualize anomalies suitable for interpreting cyberattacks. In detecting web attacks, for example, the dimensions taken from the HTTP request header were the protocol, userAgent, acceptEncoding, acceptCharset, and connection. SOMs were employed to visualize such multidimensional data to a 2D map, employing colours to distinguish anomalous web traffic. Similarly, SOMs were employed to detect botnets by reducing 5D data (i.e., protocol, source/destination IP, source/destination port numbers) to a 2D map, effectively classifying botnets from normal traffic on the map.

G. Biologically Inspired Techniques

Cyberintrusions may come not only from network traffic, but also from offending human language such as profanity, insults, hate speech, and racist/sexist remarks. To distinguish offending language from normal, Natural Language Processing (NLP) applications have emerged. NLP derives semantics from language structures such as the use of punctuation, sentence length, or a group of words frequently found together in a sentence. This allows NLP to detect sentiments, by identifying groups of words that are different from those labelled as normal.

Many biologically inspired and evolutionary algorithms are suitable to detect offending human languages. The most popular algorithm is Deep Neural Networks (DNNs), a derivative of ANNs. DNNs employ multiple hidden layers, allowing algorithms to process latent variables that are otherwise unrecognized when only one layer is used. These are suitable for NLP applications, because they can learn from language structures to derive semantics. DNNs allowed the labelling of words with their role in the sentence (e.g., adjective, noun, verb, or conjunction), finding phrases (noun phrases and verb phrases), and recognizing named entities (i.e., persons, companies, and locations).

Generative Adversarial Networks (GANs) are also a derivative of ANNs. The techniques seek to find features from data samples, given their classes. GANs consist of two sets of neural networks: one is used to generate features and the other is used to evaluate how features model the data. Their applications to cybersecurity include detecting steganography, where one set of neurons generated samples of fake images, and the other set of neurons distinguished the generated fake images from real ones. The two

sets of neurons compete against each other to reach their goal of either generating undetectable fake images, or successfully distinguishing fake ones from real, while updating their weights in each iteration.

Overall, in this section, we showed how AI techniques could improve cybersecurity solutions. The current trend shows that machine learning techniques seem to be the most popular AI-based solutions, especially when it comes to detecting network intrusions. However, as cyberattacks become more sophisticated and complex, the efficacy and efficiency of other AI-based solutions discussed here must be further explored to better evaluate their true potential in the field of cybersecurity. In the next section, we discuss how AI could be deployed in various application domains to bolster their cybersecurity posture.

SECTION IV.

Applying AI to Strengthen Cybersecurity for Various Application Domains

The Internet continues to evolve in terms of the number of users, its size, heterogeneity of devices, and the number and type of applications that are being developed to run over the internet. Today, like electricity, water, and gas, the Internet has become an important utility in the daily lives of people around the world. As more devices connect to the Internet, they face increasing risks of being exposed to all kinds of cyberattacks. To protect these Internet-connected devices along with their users, cybersecurity has become indispensable. Fig. 6 illustrates the role of AI in assisting cybersecurity in three areas namely, the Internet (section IV-A to IV-D), Internet of Things (IoT; section IV-E to IV-G), and critical infrastructure (section IV-H). The figure also illustrates the structure for the following discussions in this section: AI applications grow from two main drivers–the degree of interconnectedness, and the demand for having secure systems.

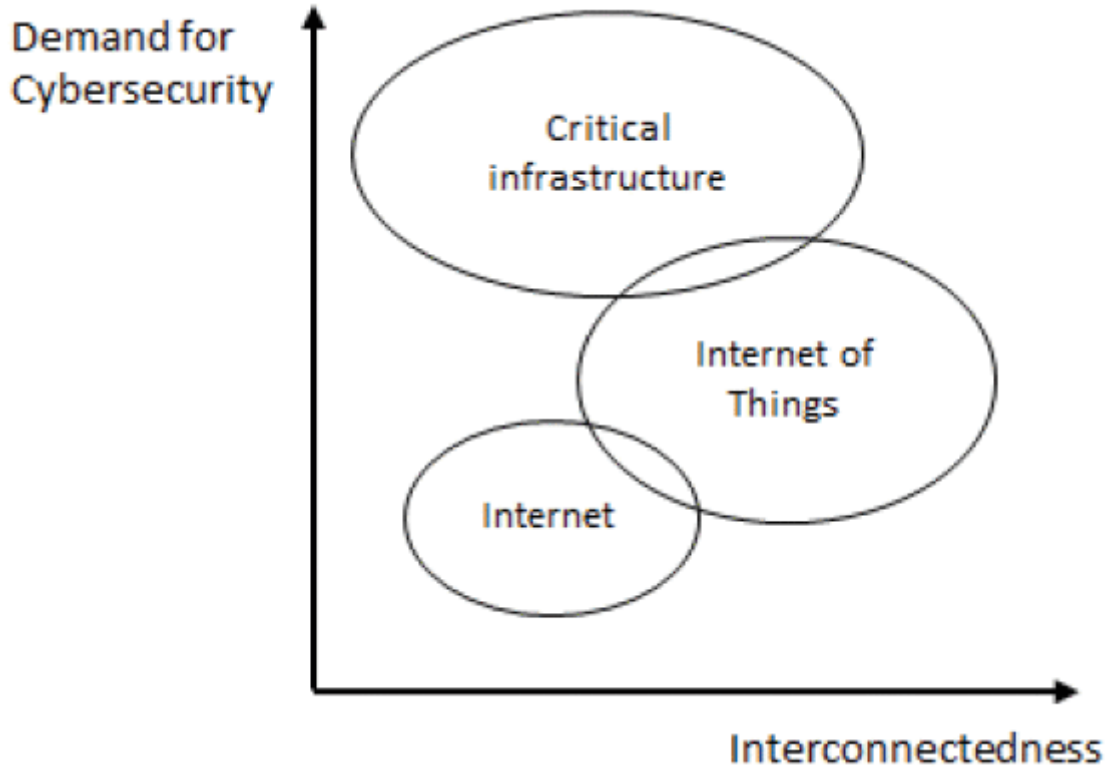

FIGURE 6. Applying AI to cybersecurity in various application domains. Larger bubble sizes reflect the heightened role of AI.

A. The Internet

From an AI perspective, cyberattacks are malicious patterns that differ from legitimate Internet traffic. To distinguish malicious traffic from legitimate traffic, intrusion-detection systems have been developed by employing AI techniques because of their capability to examine a large amount of data and adapt to the changing nature of Internet traffic. Recent cyberattacks have targeted network infrastructure, business logic, and users.

B. Network Infrastructure (Botnet)

Most Internet services involve client-server communications. Attackers can pre-empt access to servers or prevent the server from serving client requests, as in DoS attacks. In a botnet, the attackers first compromise several hosts (using Trojans or other types of malwares), which the attacker then controls and issues specific requests to execute tasks. For instance, in a DoS attack, these compromised machines can be used to overwhelm a server with many requests, leaving no resources to handle requests from legitimate users.

DoS attacks have become an increasingly serious threat as the botnets they use grow in complexity and run on multiple platforms from computers, mobile devices, and IoT devices. One study detected DoS attacks launched by IoT devices by employing features suitable to characterize IoT network behaviours. They observed that IoT devices communicate with a limited number of endpoints when running applications, so two features were proposed to reflect this:

a) the number of distinct destination IP addresses, and

b) the number of distinct IP addresses within a 10-second window.

Other features proposed were interpacket arrivals, and the first and second derivatives of interpacket arrivals. This reflects a sudden influx of packets sent by the IoT device. The study showed that decision trees achieved 99 percent accuracy in detection. Since most IoT devices must pass a single gateway (such as a home router), DoS attacks generated from IoT devices can be prevented when gateways adopt the proposed detection method.

New DoS attacks techniques are launched as new services emerge. Recent examples include DoS attacks on smart meters. Each of these meters also act as a router in the meshed network of smart meters. In, the authors found that injecting an attack packet to a meter could generate a high volume of route packets, updating other meters to change their routing information in a way that prevents data packets from reaching their destination. As such, the meters in the network exhaustively attempted to get the data packet to reach the destination, which caused the network to become unavailable. In, the authors observed that the wireless modules of smart meters are vulnerable to a jamming attack. To detect a jamming attack, they analysed the distribution of distance of the incoming wireless signal to a point calculated as central to the network. As new services and computing platforms emerge, we expect new, more complex DoS attack techniques will emerge.

Recent studies focused on detecting DoS attacks within the Software-Defined Network (SDN) environment. Network management through SDN differs from traditional forwarding protocols. While traditional routers forward traffic according to their routing tables, SDN collects and programmatically analyses network data before forwarding network traffic. This makes DoS attack detection in an SDN environment a novel challenge. The work in constructed 68 features derived from packets that an SDN system switched from its data plane, before the system forwarded packets to the control plane. These features were extracted from statistics (the ratio, entropy, count, size, and flow of packets) of the Internet

Protocol (IP), Transmission Control Protocol (TCP), User Datagram Protocol (UDP), and Internet Control Message Protocol (ICMP) packets and flags. With Deep Learning algorithms, the work showed that it detected DoS attacks with 95.65 percent accuracy.

Deep Learning is seen as a suitable solution for detecting DoS attacks in an SDN environment. The authors of employed 20 features, such as the protocol, port, and packet size, and so on. The authors showed that a derivative of Deep Learning called Long Short-Term Memory can detect DoS attacks with 99.88 percent accuracy. The work in employed a set of features, such as the number of connections within a 2-second window, duration of connections, number of connections to the same service (as the current connection), protocol type, and amount of data flow in each direction. It showed that DNNs excelled in other AI techniques, such as SVMs, Naïve Bayes, and Decision Trees in terms of accuracy. The work showed that DNNs performed well, although only a small number of features were defined, because DNNs were able to create hidden/latent variables that were considered as additional features, as opposed to other machine learning techniques that do not create features.

SDN employs AI techniques to adapt to changes in the computing environment and learn from past network data to analyse new traffic patterns and predict security trends. However, two limitations have not been addressed in the literature when AI is used for detecting cyberattacks on SDNs. First, how AI can be used for real-time detections has not been discussed. Detecting DoS attacks requires real-time decision making to classify malicious and legitimate traffic, but the solution provided through AI techniques are evolutionary in nature, which requires several computing iterations to generate the appropriate output. Although the work in tested how the proposed system performed in real time, the test was done after a classification model was obtained from training data. To the best of our knowledge, no study has proposed an AI technique for SDN to detect DoS attacks in real time. Second, SDN by its nature does not address detecting application-layer attacks. Detecting DoS attacks on application-layer protocols require either deep-packet inspection or other non-centralized techniques. This is another opportunity where AI could be applied, as we discuss in the next section.

C. Application Layer

As servers run the crucial business applications of an organization, attacking servers is an attractive venue to assault either the organization running services or their users. Until recently, application-layer attacks have focused on protocols such as HTTP, Domain Name Service (DNS), or Session Initiation Protocol (SIP). For example, when the new version of the web browsing communications protocol HTTP/2 was introduced, novel DoS attack modelling and detection was proposed in; the authors demonstrated how to bypass intrusion detection systems. HTTP/2 had a flow-control mechanism at the application layer, which did not exist in HTTP/1.1. Flooding a type of the flow controls pre-empted a server running HTTP/2 services, while maintaining a low number of connections to the target server. This bypassed known detection systems, which regard network events showing high numbers of connections as attacks. When the proposed HTTP/2 flood traffic was launched against an HTTP/2 service, AI techniques (Naïve Bayes, Decision Trees, and Rule Learning) showed a higher percentage of false alarms than when the same AI techniques were employed to detect HTTP/1.1 DDoS attacks, which demonstrated that they bypassed known intrusion-detection systems. In detecting attacks, SVMs showed no false alarms, given a proposed set of features relevant to HTTP/2 detection.

The current application-layer attack landscape has shifted from preventing information flow to manipulating information's meaning. With the advent of online social networks, a new breed of

cyberattack has emerged that aims to disseminate false information so that recipients behave or make decisions according to what the adversary intended. Probably the most influential false information was when fake news influenced the 2016 US presidential campaign, thereby affecting national security interests. False information can affect individuals, too, because it manifests itself not only in terms of fake news, but also in cyberbullying and online grooming to control the victim's behaviour. False information can seriously affect both national security and people's wellbeing; and detecting false information has become a modern application-layer cybersecurity issue.

AI has proven to be a versatile technique to detect false information, as it can quickly analyse a large amount of data. For example, in, the authors analysed a corpus of 11,000 articles, including news from Reuters, local news, and blogs, and about 29 percent of articles of the corpus were labelled as fake. Their work classified fake news with 77.2 percent accuracy using Stochastic Gradient Descent, an iterative optimization algorithm. The authors of proposed correlation-based classifiers analysed more than 150,000 tweets and showed that the proposed classifiers performed with 47 times greater precision than when the system was not employed in classifying messages. The authors analysed 4.4 million Facebook messages and classified them into fake and legitimate ones. By employing Naïve Bayes, Decision Trees, AdaBoost, and RandomForest, fake news was separated from legitimate messages with 86.9 percent accuracy.

Fake news must be detected as early as possible. Hence, a work proposed an early fake news-detection method by employing a family of ANNs. The work measured the time and structure of the propagation path in how news spread. It employed two derivatives of ANNs, i.e., Recurrent Neural Networks (RNNs) (which resemble directed graphs) and Convolutional Neural Networks (CNNs, a derivative of DNNs with more hidden layers). The CNNs measured the time propagation of news, while the DNNs measured the structure of propagation path of news, creating a tree-like structure representing how news spread from one user to another. The work was able to detect fake news in social media with 85 percent accuracy on Twitter and 92 percent on Sina Weibo within 5 minutes of when the first fake news was posted.

Furthermore, detecting false information borrows knowledge from linguistics to classify texts. Here, the text classification approaches expand observations and features required in cybersecurity to implement automatic detection methods. The features such as grammatical mistakes and choice of words are adopted from linguistic cues, which are then mapped into machine learning features. In addition, adopting specific terms with the linguistic cues, it is possible to identify bomb threats on Twitter, and identify the authenticity of Twitter users such as online predators. These works showed that automatic detection techniques for false information improve human wellbeing and demonstrate AI's capability to use new features.

In text-classification tasks, a favoured feature is tf-idf, which is short for term-frequency and inverse document frequency. The value of term-frequency increases with the number of common terms found in a document, while the value of inverse document frequency does the reverse. Many false information-detection techniques have expanded the tf-idf feature together with other linguistic cues such as phrases, grammar, negatives, and punctuation. SVMs can detect satirical sentiment in sentences that are

potentially misleading news, whereas with Naïve Bayes, it is possible to classify topics on Twitter to detect spam or phishing. DNNs have shown their ability to detect hate speech in tweets with 93 percent accuracy.

Despite recent advances in text classification tasks, detecting cyberattacks at the semantic level is still in its infancy. Studies that employed tf-idf required human intervention to supply relevant words such as "dead" or "bomb" to detect threats, and "age," "yr," or "year" to detect predators. This shows that, despite the use of AI, cyberthreat detection at the current application layer still requires human intelligence intervention. Furthermore, some studies rely on features other than linguistic cues. Examples of these non-linguistic features in detecting fake news in Twitter include the existence of URLs in tweeted messages, the ratio of followers/followees on Twitter, the number of tweets, the existence of hash tags, users' time zone, and the timestamp of when a tweet was sent. These features are specific to social media, rather than part of linguistic cues.

D. Human Link and Malware

Probably the weakest link in cybersecurity is the human who is the end user of the Internet. Humans are focused on their business tasks rather than constantly dealing with the ever-increasing number of cyberattacks. While machines can be re-engineered to mitigate some of the well-known cyberthreats, humans require constant training based on past and updated issues. This requirement is one of the main reasons behind the success of malware spreading through modern phishing techniques.

Malware is software (such as a virus, Trojan, or worm) that has malicious intent. Phishing is a method that attempts to trick human users to perform what an adversary intends to do, such as clicking a link or an executable file. Such actions either trigger the spread of malware or induce the victims to reveal their sensitive information. Traditionally, phishing techniques leverage human weaknesses in their sensory systems, such as through fake emails or websites, causing victims to be unable to distinguish them from legitimate ones. Current phishing techniques are more sophisticated in that they exploit the human limit in becoming omniscient. To avoid falling for phishing hooks, users must assess the target's legitimacy, and often this can be done by inspecting the code behind the links, which may require some specialized expertise. This is an area where AI can be used to augment human intelligence.

Instead of having to learn all the rules on how to detect phishing, these rules act as the features for AI techniques. The authors of proposed an approach that uses SVMs to detect links, leading to false banking websites. The approach uses five features: IP address, Secure Sockets Layer (SSL) certificate, number of dots in the URL, web address length, and blacklist keywords. Legitimate banking websites show a legitimate domain name instead of an IP address, have an SSL certificate, have relatively short URL lengths in the domain, and are not part of a subdomain (higher number of dots). Furthermore, the method collected a bunch of words commonly used in phishing websites. The results showed that the method was able to detect zero-day phishing with 98.86 percent accuracy. This research demonstrates that with AI training, we can address the human weaknesses in cybersecurity awareness.

Adversaries continue to exploit human weaknesses, as seen in attacks on modern websites and online social media. Modern websites improve web browsing experiences using JavaScript to increase user-browser interactivity and browser response time. Adversaries can leverage JavaScript either to insert malware or phish users. Detecting JavaScript-compromised websites requires advanced knowledge in coding, causing such compromised websites to become nearly impossible to detect by the average

human user. Furthermore, recent techniques spread malware through online social media by phishing for users to click on a link, causing users to unintentionally download malware (also referred to as drive-by-download. In response, AI techniques have been employed to detect malicious JavaScript websites and drive-by-download attacks. In this case, AI techniques have been employed to analyse JavaScript word sizes, the distribution of coding characters, frequency of bytecode in strings, commenting style, and sensitive function calls, to overcome human limitations in detecting and analysing such features. Furthermore, another approach based on AI has been used to detect an obfuscated malicious JavaScript and provide fail-safe mechanisms to prevent malware spread after users have been phished.

In the area of usable security, the goal is to create usable yet secure systems for the average human user. One approach to increase cybersecurity awareness of the average human user is by using some forms of games. The game sharpens players' vigilance in detecting fake URL forms that appear like the authentic ones; for example, distinguishing the fake URL "http:\\www.paypa1.com" from the authentic "http:\\www.paypal.com". In, the authors examined 28 papers that discuss cybersecurity training games. While the results from the examined papers revealed that the players liked the game, those papers did not show how effective the games were. Their sample sizes were small, the participants were selected (rather than randomly invited), and the effect size (i.e., the difference in cyber awareness between the group that played the game and a control group) was not studied. Furthermore, critics argue that such training games suffer from privacy and trust issues. Such training games require algorithms to learn about users' belief in their own ability to accomplish a certain goal, their attitudes toward software updates, creating strong passwords, identifying potentially malicious links, and using appropriate hardware (e.g., backup data). When information learned from the algorithms went into the hands of an adversary, the information would become useful ingredients to create tailored phishing attacks toward a target. The escalated issue would be when if such data becomes public or available to unauthorized parties, leading to privacy and trust issues.

E. The Internet of Things

Computers have become smaller, portable, and more powerful and affordable. The ubiquity of mobile devices such as phones and tablets became the dawn of the IoT era. Today, many devices (from toys, appliances, and vehicles to industrial control systems) are equipped with networking capabilities and Internet connectivity that makes the IoT possible. Fig. 7 illustrates the evolution of technologies that have led to the emergence of the IoT. Other paradigms such as cloud computing, big data, and fog computing are enabling mobile devices with limited resources to access a wide range of services remotely.

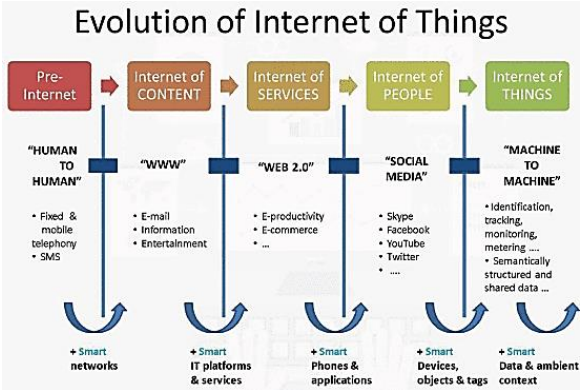

FIGURE 7. From Internet of Content to Internet of Things (Short Message Service [SMS])

Since the demand for higher data rates keeps increasing, researchers introduced fog computing services by provisioning the platform and application closer to the user. Fog computing distributes servers to minimize network roundtrip delays, especially for Content Delivery Networks (CDNs). So, fog computing improves website performance, and provides real-time energy and carbon footprint management. Furthermore, advances in telecommunications technologies led to the development of vehicular networking applications, which enable fast data transfers between mobile devices.

F. Privacy

As Internet-connected devices become smaller and pervasive, their ability to capture data surpasses humans' ability to become aware of their activities (in capturing data). Devices collect information such as voice, geolocation, surrounding temperature, and ambient illumination to improve user experience. However, studies show that collecting such information can serve malicious intent. Intelligent virtual assistants (such as Amazon Alexa, Apple's Siri, and Google Home) can be used to illegitimately open a smart (garage) door or record private conversations. One study showed that devices can be used to find a place in an airport to smuggle, cyberbully, spread fear, and divert one's browsing journey to serve advertisements. Devices also can be used to tag a location or person with crime-related incidents.

Traditionally, privacy has been addressed through secure authentication mechanisms, such as encryption and security certificates. These mechanisms shift in the IoT, as devices are mobile, with data stored in the cloud. AI techniques can be used to maintain private communications when routing paths dynamically change, and when a third party stores the data. For example, learning automata was adopted to distribute secure certificates to moving vehicles, and artificial immune system algorithms were adopted to securely self-organize Wireless Sensor Network (WSN) ad hoc connections to serve mobile gadgets. In WSNs, different IoT devices such as mobile gadgets dynamically join and leave the network. This causes traditional security measures such as port security (i.e., restricting traffic only to a known Media Access Control (MAC) address) inapplicable. Thus, the authors proposed features such as packet receiving rate, packet mismatch rate, and energy consumption per packet received from a device to describe a device's behaviour. They used artificial immune system algorithms to classify a device's behaviour as normal/abnormal. Upon detecting abnormal behaviour, unencrypted packets were dropped. This shows why an increasing number of Internet-connected devices require new privacy solutions. Furthermore, because substantial amounts of data are stored in the cloud, privacy concerns arise in relation to how sensitive data can be accessed by cloud operators. To address this issue, intelligent algorithms were employed to distribute sensitive data into several cloud servers, making it impractical for cloud operators to eavesdrop.

Secure authentication mechanisms also made use of well-known biometrics and human behaviour metrics. However, issues arise when authentication devices cannot find a good fit under varying operating conditions. To address these issues, AI techniques (such as Genetic Algorithms) have been used to enable robust performance and accurate detection of face, fingerprint, and voice recognition in different operating environments.

One disruptive technology that can bypass legislation to promote privacy is blockchain. Blockchain allows a network of peer-to-peer non-trusted computers to store encrypted data without a central authority's involvement. AI techniques are used in conjunction with blockchain to facilitate blockchain applications. In, AI techniques enable blockchain applications to guarantee secure communications between two IoT devices. Security measures that allow two IoT devices to remotely communicate have

traditionally been based on some centralized systems. Thus, blockchain was proposed to allow a pair of remote IoT devices to communicate securely without using a centralized system. Information obtained from Reinforcement Learning stored in the blockchain was used to assess whether the communicated data fulfils the end devices' access control policies, allowing automatic resource sharing between IoT devices.

The work in described how the healthcare sector could derive medical data for predicting potential diseases or medical issues while respecting patients' privacy. Classification and prediction algorithms require substantial data, which conflicts with the patients' interest in sharing their medical data. Blockchain could be employed to record such medical data, allowing patients guaranteed privacy while enabling them to take control of their personal data, such as managing access privileges. By having a platform that protects their privacy, patients have more trust in storing personal data and biomarkers (e.g., blood parameters, waist circumference) useful for providing health status and risks. AI techniques such as DNNs could be used to derive features such as biomarkers and tumour tissues from medical imaging data before being recorded to the blockchain. RNNs could be used to identify chronic conditions and predict potential diseases (e.g. cardiovascular or diabetes) from medical records.

AI techniques such as similarity learning were employed in a smart, contract-based, data-trading system. But a controversy arises when the data downloaded by the purchaser is not consistent with what the provider claimed. Thus, similarity learning was employed to calculate the distance between the purchaser's and provider's data features, thereby verifying the data's consistency. This shows that AI roles in privacy will incorporate legal, regulatory, and ethical frameworks, as sharing personal data can benefit human wellbeing.

G. Cyber-Physical Systems

Cyber-Physical System (CPSs) integrate communication, computation, and monitoring functions. They collect data using sensor networks and embedded systems and respond to the environment through software components and actuators. The fundamental CPS concepts are being deployed worldwide, as countries compete to become a dominant player in this domain. The phenomena described in CPS are behind the motivation for the economic development in Germany's "Industry 4.0", China's "Made in China 2025", and western countries' "Smart City", where manufacturing processes are automated, and suppliers at different locations link to each other. CPS may be viewed as the new AI-driven economy.

One of the earliest requirements that motivated intelligent manufacturing was to develop products within a shorter time. AI techniques were employed to autonomously collect data and collaboratively accomplish tasks to produce electronic circuit boards, control systems to perform real-time analysis on remote hydroelectric power plants and assess reliability and safety on railway control systems. Another major driver behind employing AI in intelligent manufacturing was the education sector, which requires adaptability to individual learners. To meet this requirement, educational software using intelligent agents was developed, to adapt to students' learning pace by adjusting levels of difficulty on presented exercises.

AI techniques are suitable to address the requirements of CPS, because they yield accurate predictions and estimates of outputs. The energy management sector was among the early adopters of AI techniques, to predict temperature given the changing environment. In this case, fuzzy networks were used to control air conditions for the desired temperature output. On a larger scale, power distributions demand improved energy quality, capacity, and reliability. AI techniques such as genetic algorithms and neural

networks have also been adopted in this area. They are used to solve profit management problems, where selling and buying to/from the grid are subject to varying energy tariffs.

The need for CPS stems from the ubiquity of small devices, which enhances the capability to collect data, thereby providing the opportunity to process big data. This is an area where AI applications in CPS converge with AI applications in cybersecurity, because often data is remotely collected via processing systems. In this case, cybersecurity issues include how to collect data with a high level of trust, transmit it securely, and share it while preserving the data's integrity and privacy. The AI applications in CPS converge with previous discussions on secure networks, reliable data, and privacy issues.

AI applications' convergence in CPS with cybersecurity is readily apparent in smart agriculture, where sensors are installed in the soil to collect temperature information and levels of nitrogen and carbon. Farmers combine their sensor data with real-time data of weather predictions to make informed decisions in utilizing water and fertilizer to develop an irrigation-monitoring system. The system is employed in AI techniques, using genetic algorithms to calculate the threshold for an acceptable temperature. Sensor-based systems use cloud applications to store and process the various sensors' data, thereby providing farmers with real-time data. This allows farmers to reach optimum crop-production quality. Cybersecurity issues arise if any of these cyber entities can be attacked–from sensor-infecting malware, the integrity of data transmitted through the network, and the availability of cloud computing resources to the irrigation system, to whether sensor data can be shared. Failure to address such cyber issues can seriously affect crop harvesting.

H. Critical Infrastructure

Critical infrastructures are assets that fundamentally support national security and society. These infrastructures include power (oil, gas, electricity, and nuclear), water, air traffic control systems, and telecommunications. Thus, safeguarding critical infrastructures are of paramount importance because people's daily activities and lives depend on their availability and integrity. Previous discussions showed how cybersecurity has expanded in its scope from network intrusion detection systems to how human wellbeing could be improved. The shift was motivated by different sectors, such as health and education. Additionally, the critical infrastructure sector also fuels the development of AI techniques to enhance cybersecurity.

Cybersecurity's role in critical infrastructures is mainly associated with securing SCADA systems. They are the main infrastructure's control systems (consisting of computing nodes that communicate with other nodes). SCADA systems typically reside on Operational Technology (OT) networks of the organization. As these OT networks and Information Technology (IT) networks become more closely intertwined and connected to the Internet, they are increasingly vulnerable to external and internal cyberattacks.

Despite these risks and their inherent vulnerabilities, critical infrastructures must be resilient against such cyberattacks. Hence, one of the requirements and challenges is to maintain a critical infrastructure's business continuity. Maintaining the SCADA systems' resiliency can be accomplished by applying AI techniques. For example, in wind turbine generators, faults could be predicted by employing Artificial Neural Networks (ANNs) that monitor ambient temperature, generator speed, and pitch angle of the generator power outputs. In controlling water systems, AI techniques such as k-NN, Decision Trees, and SVMs were employed to classify different anomaly events, including cyberattacks and hardware failures. Furthermore, AI techniques such as SVMs and ANNs have been used to provide access control

to SCADA systems based on users' dynamic attributes, such as location, time of use, and the user's work shift (when the user works onsite). Using AI to build robust resiliency will remain an active research area, because of the high importance of the critical infrastructure sector in society.

Other AI techniques, such as propositional logic, have been adopted around critical infrastructure protection. The authors proposed a logic-based framework to enforce security policies for system authorization in SCADA systems, because the authentication process in this environment requires complex mapping between user privileges and system rules. In such a framework, rules are distributed across system nodes, so that they can derive the sets of actions the user can perform on each node. When a user with a certain privilege sends a command to a destination node, both the user privilege information and the command are sent to an authorization server. The server analyses the information received, generates a token, and forwards all the information (i.e., user privilege, command, and token) to the destination node. The node analyses the token with its local authorization policy, to allow/disallow the command's execution. Thus, the proposed logic-based framework promotes scalable authentication in SCADA systems, because the authorization decision of allowing/disallowing commands takes place at destination nodes.

Intelligent algorithms employing logic have also been proposed to self-heal SCADA systems' communications channel. SCADA systems secure their communication with remote nodes using session keys. In the event of a node failure, it is critical for the node to immediately re-establish the communications channel before any unauthorized user/agent takes control over the re-establishment of the communications channel. Thus, the authors proposed distributing re-keying materials to the remote nodes, which is required to generate a new session key. The re-keying materials consist of a series of numbers generated from a mathematical formula (i.e., bivariate polynomial). Similarly, generating a session key goes through mathematical and logic processes to generate a session key. Thus, after a remote node is recovered from an unavailability incident on its communication channel, the node can generate a session key, effectively self-healing the communications channel.

Furthermore, mathematical models also have been used to self-heal electrical distribution systems upon encountering faults. After such events, the self-healing system determines which network zone to isolate based on a set of 22 features such as the cost of power losses, power demand at each node, and the voltage magnitude at each node. The system employed set theory to cluster the features. Afterward, the system fed these clusters to a series of mathematical models (i.e., backward/forward sweep load-flow algorithms) that represent the steady-state of electrical distribution systems. Thus, both logic and mathematical methods are being widely used to meet the cybersecurity requirements of the critical infrastructure sector.

Table 2 summarizes the discussion results of this section. As the Internet evolves, the role of AI in cybersecurity will broaden. AI techniques are being employed in applications that are critical to national security and human wellbeing. Not only are AI approaches being used to solve problems rationally, but also to make machines think and act like humans.

Domain area		Cyberattacks	Challenges	AI solutions
Internet	Network	Denial of Service	Changing traffic patterns; large number of features	Learn changing traffic patterns; increase accuracy with a small number of features
	Application layer	Changing the semantics of messages; fake news	Big data; specific features from linguistics	Classify semantics based on grammar, choice of words, negatives, sentiment, user authenticity
	Human	Phishing	Training people is difficult; changing and varying attack methods	Automatic phishing detection, malicious links, malicious JavaScript
Internet of Things	Privacy	Information assurance; impersonation	Whether data can be shared or must be secured	Secure data in distributed environment
	Cyber-Physical Systems	Insecure data collection and sharing (e.g. cloud)	Large, distributed area	Benefit management; cloud security
Critical infrastructure		All attacks in the cyber-attack landscape	Build resiliency	Logic-based framework

SECTION V.

Future Challenges and Research Opportunities

A. The Race Between Défense, Offense, and Humanity

Recent AI research advances in cybersecurity have fuelled the race between the white hat (defenders) and black hat (offenders) hackers. Attackers can employ AI to mimic human behaviour to achieve personal pride, power, or financial advantage. AI has led to the creation of intelligent agents that automatically click advertisements, play online games, and buy and resell best-seller seats for concerts. AI has also manipulated public opinion in Venezuela by retweeting political content and has affected the US presidential election by spreading tailored news. Future research opportunities in cybersecurity are determined by how dividing lines can be drawn between developments and basic needs.

AI's use in cybersecurity impacts three major stakeholders: white hat hackers, black hat hackers, and end users (humanity). The white hat and black hat hackers are the cohorts who promote the development of AI techniques. However, it is difficult to find the dividing line between the two groups to regulate technological deployment, because one's advancement follows the other's advances. Hence, it is imperative to investigate how AI can be employed for human basic needs and for developing cybersecurity controls.

B. Infrastructure

The use of AI in cybersecurity is viewed as a race between law enforcement and cyber attackers. The leader in the race will be determined by his/her access to technical knowledge and the supporting computing infrastructure. AI algorithms are computationally expensive because they are evolutionary by nature. For example, to detect malware, hashing algorithms have been developed to input to the k-means clustering algorithms, to enable fast clustering of common data samples. Developing relevant algorithms has become part of the recent race, but hardware development is another crucial part.

C. Hardware and Platform

Having access to state-of-the-art computing infrastructure will help solve AI problems efficiently and with efficacy. As the number of computing devices increases, the volume of traffic will also increase, thereby making it necessary to perform data analysis quickly. Consequently, analysing data by using AI techniques requires high-end computing platforms. To address this challenge, cluster computing solutions such as Apache Spark and Hadoop have been employed to analyse cyber traffic. At the high end, quantum computing will be the breakthrough technology that helps solve complex computing problems. NASA's quantum computer has been able to solve complex problems in a fraction of time–it is 100 million times faster than traditional computers.

D. Resources

Having easy access to the required resources when needed is crucial in implementing workable computing solutions. Currently, energy is seen as the scarce resource for many computing needs. For instance, Bitcoin blockchain consumes an equivalent energy of 29 average Australian households for a full day, only to commit one block.

When intelligent computers start to consume a significantly larger chunk of resources which are shared with human beings, ethical issues regarding the use of AI will arise. One issue would be if intelligent machines have their own rights. In one way, the issue may seem irrelevant because computers are viewed as having no consciousness. In another way, researchers have started to debate whether intelligent computers should have rights regardless of the definition of consciousness. The adoption of AI in cybersecurity extends the arguments on how to share scarce resources between intelligent computers and human. This will in turn motivate regulators to go back to the drawing board to justify what serves as development and basic needs. Ethical issues will also remain a future challenge when it comes to how AI can be employed for cybersecurity.

SECTION VI.

Conclusion

As the speed and sophistication of attacks increase, AI has become an indispensable technology in the cybersecurity area. This article showed how cyberthreats have increased, evolved in their complexities, and broadened their scope. We underscored how past cyberthreats remain relevant to future risks. We presented a comprehensive review of cyberthreats and solutions. We described how cyberattacks can be launched on different network stacks and applications, along with their impact. Cyberthreats will continue to rise, even as the community identifies cyberthreats and develops solutions using a wide range of technologies and techniques.

In contemporary research, AI techniques have demonstrated their promise in combating future cybersecurity threats. The techniques propose a range of intelligent behaviours—from how machines can think to act humanly. Recently proposed AI-based cybersecurity solutions largely focused on machine learning techniques that involve the use of intelligent agents to distinguish between attack traffic and legitimate traffic. In this case, intelligent agents act as humans whose task is to find the most efficient classification rules. However, the cyberattack landscape today morphs from disrupting computers to sowing disorder in society and disturbing human wellbeing. We discussed this phenomenon in terms of how advances in technologies are transforming the ways cyberattacks can be launched, detected, and mitigated. Through such advances, AI's role in cybersecurity will increase continuously. Novel AI techniques must be developed to quickly detect and mitigate threats that impend upon societal and human wellbeing. Likely, cybersecurity solutions will expand from intelligent agents acting humanly to thinking humanly.

Although AI's role in solving cybersecurity issues continues to be investigated, some fundamental concerns exist surrounding where AI deployment can become regulated. For instance, as intelligent machines become more integral solutions for humanity, these machines increasingly will consume fundamental resources for life. When humans and machines compete for scarce resources, a new form of governance will promulgate. This in turn will engender a new research avenue.s