IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 5, 2020, accepted January 14, 2020, date of publication January 20, 2020, date of current version February 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2968045

INVITED PAPER

Harnessing Artificial Intelligence Capabilities to Improve Cybersecurity

SHERALI ZEADALLY[1], ERWIN ADI[2], ZUBAIR BAIG[3], AND IMRAN A. KHAN[3]

[1]College of Communication and Information, University of Kentucky, Lexington, KY 40506-0224, USA
[2]UNSW Canberra Cyber, University of New South Wales, Northcott Drive, Campbell, ACT 2600, Australia
[3]School of Information Technology, Deakin University, Geelong, VIC 3220, Australia

Corresponding author: Sherali Zeadally (szeadally@uky.edu)

ABSTRACT Cybersecurity is a fast-evolving discipline that is always in the news over the last decade, as the number of threats rises and cybercriminals constantly endeavor 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. Traditional cybersecurity solutions are becoming inadequate at detecting and mitigating emerging cyberattacks. Advances in cryptographic and Artificial Intelligence (AI) techniques (in particular, machine learning and deep learning) show promise in enabling cybersecurity experts to counter the ever-evolving threat posed by adversaries. 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.

INDEX TERMS Artificial intelligence, cybersecurity, cyberattacks, machine learning.

I. INTRODUCTION

Cybersecurity is defined as a set of processes, human behavior, and systems that help safeguard electronic resources. Analogous 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 [1]. Global cybersecurity spending is expected to exceed $1 trillion from 2017 to 2021 [2], where spending on cybersecurity already increased by almost 40 percent from 2013 to $66 billion [3].

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 cyberattacks while hiding their tracks. However, in this work, we focus on how AI-based cybersecurity solutions could fend off attackers better, and minimize or prevent data breaches.

The associate editor coordinating the review of this manuscript and approving it for publication was Chi-Yuan Chen.

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 [4]. Today, AI has been deployed in far-reaching application areas, including healthcare, agriculture, space, law, and manufacturing [5]–[9]. 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. A decade ago, a subfield of machine learning, also known as deep learning, emerged that enables machines to discover hidden relationships in their input data, thereby generating more accurate results for planning and predicting. 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 analyze 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 [10]. The key sources of cyberthreats include [11]:

1) *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.
2) *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 as a result of data breaches, seek payments through ransomware, and so on.
3) *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.
4) *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.
5) *Spies (including business rivals)*: They steal trade secrets to gain an unfair market advantage.
6) *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.
7) *External attackers and insider threats*: Experts with a strong knowledge about the operation of computing resources as well as human behavior, 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 antivirus. 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 [12].

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 malware every minute [13].

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

1) *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 [1].
2) *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 behaviors, data mining focuses on the discovery of previously unseen patterns in data (very much similar to 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 [14]–[16].

3) *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 [17]. 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 cyberprevention 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 [18].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 [19]. Precedence is typically given to safety, reliability, robustness, and maintainability of such systems, and therefore security takes a backseat [20].

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.

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 [21], the top 10 cyberthreats we face today include:

1) *Denial of Service (DoS) attacks*: These attempt to overwhelm a victim system's computing resources by sending an overwhelming number of requests for it to process within a short period of time. Such attacks can be carried out in one of several ways: a single attacker machine can launch a DoS attack against a victim machine by transmitting a large number of network traffic packets that appear to be legitimate, to bypass security controls along the way; multiple attacker machines can participate in a distributed-style DoS attack, i.e., a Distributed Denial of Service (DDoS) attack, resulting in a similar outcome at the victim machine. DoS attacks are increasingly becoming more sophisticated and harder to detect, because of the ready availability of attacker tools, as well as the proliferation of the CyberCrime as a Service (CCaaS) market [22].
2) *Man-in-The-Middle (MiTM) attacks*: These are legacy cyberattacks carried out through the process of interception of transmitted data on a communication line between two legitimate communicating parties. The attacker places itself either physically or virtually between two communicating parties, A and B, posing as A to communicate with B through the interception of A B messages and replacing these with malicious or tampered messages, and repeating the same process on the BA communication line, i.e., posing as party B and speaking to party A. Variant implementations of such an attack include IP address spoofing, wherein the malicious actor convinces legitimate systems that it is a trusted entity, enabling system access for the actor. A message replay attack involves the repeat transmission of a previously stored, stale message on the communication line, perpetrated by the malicious actor.
3) *Phishing and spear-phishing attacks*: These are carried out by crafting emails that appear legitimate and transmitting them to legitimate systems, with the intent of having the naïve end users click a link and divulge personal information. Such attacks exploit social engineering principles, wherein emails are made to appear legitimate to end users, luring them to trust them. Spear phishing is defined as a carefully designed attack that involves a thorough background search carried out by the malicious actor on susceptible victims, for subsequent drafting of emails that appear to be very legitimate, with the ''from'' field often containing trusted email addresses.
4) *Drive-by attacks*: These are carried out by malicious actors who skim through the web and search for vulnerable websites, so that they can implant malware scripts into the webservers. End users who visit the website are eventually infected with the malware, leading to system compromise, disclosure of sensitive data, and other damage.
5) *Password attacks*: These can be carried out by shoulder surfing user keyboard activity, brute force into a system using common passwords, and crafting sophisticated passwords through the application of AI techniques [23], [24].

6) *Structured Query Language (SQL) injection attacks*: These are legacy cyberattacks that exploit vulnerabilities in the SQL language by injecting a webpage with input fields with SQL query code, that when executed at the webserver, would disclose some or all of the stored content on a backend database server, possibly including usernames and passwords.
7) *Cross-site scripting attacks*: These are carried out by injecting malicious code in a vulnerable webserver. Subsequent retrieval of the hosted webpages by naïve end-users would infect the victim's machine with malware. Such malware may transmit user data from the victim's machine to the malicious actor's servers, and may lead to the subsequent hijacking of web sessions, theft of credentials, installation of key stroke loggers, capture screenshots, and even taking control of the victim's machine remotely.
8) *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.
9) *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 [25]
10) *Malware attacks*: One of the main difficulties to web-hosting organizations is that their websites can become the source of malware spread. According to Symantec's 2016 threat report, 78 percent of websites contain a critical vulnerability that can be exploited by the adversary to allow malicious code to run without any user interaction [26]. Strengthening a website's defenses involves deploying appropriate security controls such as web proxies, firewalls, and intrusion detection systems. A major issue here is the tradeoff between the right level of security controls and usability of websites being hosted. The higher the level of a website's usability, the greater the area of vulnerability for the website.

Network attacks 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.

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

1) *Game theory:* This has been previously applied to cybersecurity [27]–[29]. 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 behaviors 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 [30]–[32]. At each step of the game, the attacker and the defender would adopt strategies to be successful in their respective goals [33].
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 [34].
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 [35], performs a sequence of steps comprising substring matching in order to identify suspicious website addresses. The second phase of the presented scheme comprises the scanning of the web address through the VirusTotal application (i.e. a website where one can supply a web address and gets a scored analysis about the degree of maliciousness of the input website), 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 [36]. 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

TABLE 1. Various types of attacks, their impact, and approaches to detect them.

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

policies [37], mathematical models [38], and statistical techniques [39]. 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 [40], 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 defense against DoS/DDoS attacks by establishing a communication path between multiple cyberneurons deployed on different devices. Last, Central Nerve serves as a knowledge base against new attacks and to disseminate information to other security devices. Collaborative defense 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 [41]. End users might not update their devices with the latest security patches, with some vendors attempting to push automatic updates, in order to install security patches. The Wannacry ransomware [42], [43] 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 [41] 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.

III. ARTIFICIAL INTELLIGENCE

AI is concerned with how machines can think or act correctly, given what they know [44]. This universal definition includes how closely machines can think or act like humans (Fig. 1). 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 [45] sets the standard on machine intelligence. Under

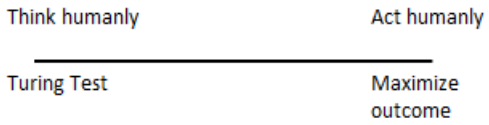

FIGURE 1. Spectrum on intelligent measures from thinking humanly through the Turing Test, to acting humanly to maximize the outcome.

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 [46], checker games [47], and a family of blocks-world problems.

After the proliferation of the Internet in the late 1990s, 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 vandalism in Wikipedia articles where anybody can contribute as authors [48]. In contrast, malicious bots also emerged to cheat in online games [49], post spams [50], [51] and spread malware [52]. In mimicking online games, bot programmers analyzed the traffic flow between the game console and server to reverse engineer the game code [49], [53]. In posting spams, the bots mimicked the behavior of human when online, such as surfing the pages before posting a message in a forum, rather than continuously posting messages [51]. 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 [49]. Furthermore, game bots are more interested to collect items, while human players seek to collaborate with other players to complete challenges/quests [48]. Similarly, spambots and malware bots can be detected from their behaviors being different than human, that can be detected through some distinctive communication patterns [50], [52].

The most relevant AI applications to the cybersecurity area are in intrusion detection systems [54]. 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 analyze this traffic have become time-consuming and make security protection systems behave defensively rather than proactively. Coupled with this trend, technological advances are also benefiting attackers who are developing new sophisticated attack strategies that can avoid detection by current security systems [4]. 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 analyze 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 [55]–[57] as they can handle large volumes of data and a wide range of data attributes (e.g. a large number of table columns) used for classification [54], [55]. 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 labeled according to their class (e.g., malicious or legitimate). Training data, or data labeling 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 labeling 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 [58]. This shows that taxonomy perspectives are converging, making it less essential to define machine learning algorithms based on whether they are supervised or unsupervised [59]. Henceforth, we present an in-depth discussion of machine learning algorithms from a taxonomy perspective as described in [60], 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

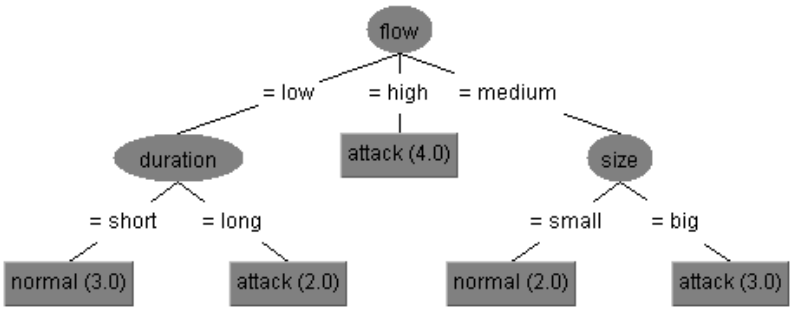

FIGURE 2. An example of a decision tree that classifies network traffic into attack and normal traffic type.

their features. Naïve Bayes is a machine learning technique used to classify data based on the Bayesian theorem [61] 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 [62]–[65], because it can inherently accept such a naïve assumption (that each feature comes from independent events) while still yielding acceptable results [66].

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. 2 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 [67]. 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 [68]. 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.

A different approach is the Rule-Learning technique [69], 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 similar to 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 [70]. 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 [71], [72].

C. K-NEAREST NEIGHBORS

The k-Nearest Neighbor (k-NN) technique learns from data samples to create classes or clusters. It was first proposed as a non-parametric pattern analysis [73] to find the proportion of data samples in a neighborhood that yields a consistent estimate of a probability. The neighborhood 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 neighbors decide how new data samples can be assigned to one of the clusters.

Fig. 3 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 neighboring cluster. Hence, when $k = 3$, the sample was put into Class 2. When $k = 9$, the sample was put into Class 1.

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

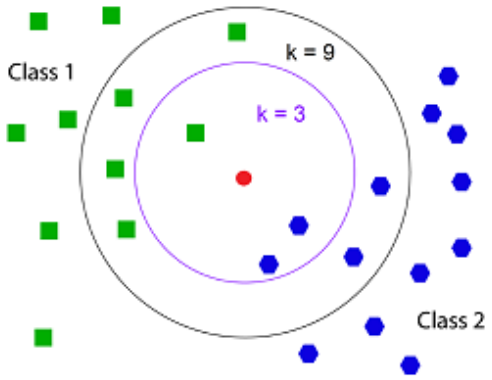

FIGURE 3. The k-Nearest Neighbor (k-NN) algorithm classifies data in class 1 and class 2, based on the k nearest data samples in the neighborhood from the new data sample.

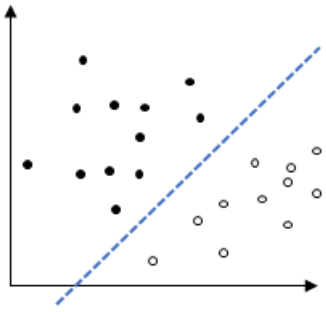

FIGURE 4. Support Vector Machines (SVMs) find a plane that separates data samples.

real-time detections of cyberattacks [74]. Recently, the technique was employed to detect attacks such as data tampering and false data injection against industrial control systems [75] and smart grids [76]. 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 [75] or a vector [76].

D. SUPPORT VECTOR MACHINES

The Support Vector Machines (SVMs) [77] technique extends linear regression models. While classifying data samples, SVMs find a plane that separates data samples into two classes (as shown in Fig. 4).

The separating plane can be shaped to form linear, nonlinear, polynomial, Gaussian, Radial, sigmoid, and so on depending on the function employed (called a kernel) [78]. 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 analyze 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) [79].

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 [80]. 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 [81]. 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 [82]. 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 [83], an ANN application used the Cascade Correlation Neural Network (CCNN) [84] 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 labeled training data, allowing the neurons to adjust their weights and detect unseen DoS attacks [85]. 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 labeled when an incident (such as DoS) occurred, allowing the detection system to learn from the incident.

F. SELF-ORGANIZING MAPS

Self-Organizing Maps (SOMs) [86] 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 [87]. Without visualization, the outputs from intrusion-detection systems are hard to analyze. 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 a large number of 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 [87], [88]. 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 colors to distinguish anomalous web traffic [88]. 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 [89].

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 [90], [91]. To distinguish offending language from normal, Natural Language Processing (NLP) [92] 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 labeled as normal [90].

Many biologically inspired and evolutionary algorithms [93] 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 [94]. DNNs allowed the labeling 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) [95] 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 [96], 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.

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, similar to 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. 5 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:

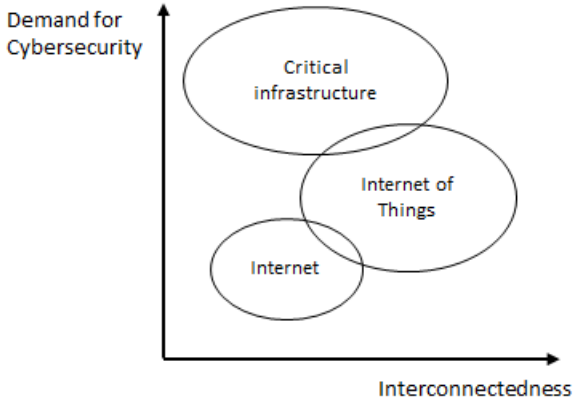

FIGURE 5. Applying AI to cybersecurity in various application domains. Larger bubble sizes reflect the heightened role of AI.

AI applications grow from two main drivers–the degree of interconnectedness, and the demand for having secure systems.

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 preempt 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 malware), 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 a large number of 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 [97] detected DoS attacks launched by IoT devices by employing features suitable to characterize IoT network behaviors. 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 [98], [99]. Each of these meters also act as a router in the meshed network of smart meters. In [98], 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 [99], the authors observed that the wireless modules of smart meters are vulnerable to a jamming attack. To detect a jamming attack, they analyzed 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 [100]–[102] 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 analyzes network data before forwarding network traffic. This makes DoS attack detection in an SDN environment a novel challenge [103]. The work in [100] 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 [101], [102]. The authors of [101] 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 [102] 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 analyze new traffic patterns and predict security trends.

However, two limitations have not been addressed in the literature [100]–[102] 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 [101] 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 [103]. 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 modeling and detection was proposed in [104]; 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 preempted 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 [105]. 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 [104].

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 [106]. Probably the most influential false information was when fake news influenced the 2016 US presidential campaign, thereby affecting national security interests [107]. 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 behavior. 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 [108]–[110], as it can quickly analyze a large amount of data. For example, in [108], the authors analyzed a corpus of 11,000 articles, including news from Reuters, local news, and blogs, and about 29 percent of articles of the corpus were labeled as fake. Their work classified fake news with 77.2 percent accuracy using Stochastic Gradient Descent, an iterative optimization algorithm. The authors of [109] proposed correlation-based classifiers, analyzed 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 of [110] analyzed 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 [111] 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 [112] to classify texts. Here, the text classification approaches [108], [109], [113], [114] 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 [109], and identify the authenticity of Twitter users such as online predators [114], [115]. 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 favored 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 [109], [113], [116] 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 [113], whereas with Naïve Bayes, it is possible to classify topics on Twitter to detect spam or phishing [117]. DNNs have shown their ability to detect hate speech in tweets with 93 percent accuracy [116].

Despite recent advances in text classification tasks, detecting cyberattacks at the semantic level is still in its infancy. Studies that employed tf-idf [109], [113]–[115], [117] required human intervention to supply relevant words such as ''dead'' or ''bomb'' to detect threats [109], and ''age,'' ''yr,'' or ''year'' to detect predators [115]. This shows that, despite the use of AI, cyberthreat detection at the current application layer still requires human intelligence intervention. Furthermore, some studies [110], [111], [115], [118], [119] rely on features other than linguistic cues. Examples of these nonlinguistic features in detecting fake news in Twitter include the existence of URLs in tweeted messages [110], [118], the ratio of followers/followees on Twitter [118], [119], the number of tweets, the existence of hash tags, users' time zone [115], and the timestamp of when a tweet was sent [111]. 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 [120].

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 [121], 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 [122], 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 [123] 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) [124]. In response, AI techniques have been employed to detect malicious JavaScript websites [125], [126] and drive-by-download attacks [127]. In this case, AI techniques have been employed to analyze 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 analyzing such features. Furthermore, another approach based on AI has been used to detect an obfuscated malicious JavaScript [128], and provide fail-safe mechanisms to prevent malware spread after users have been phished [129].

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 [130]. The game sharpens players' vigilance in detecting fake URL forms that appear similar to the authentic ones; for example, distinguishing the fake URL ''www.paypa1.com'' from the authentic ''www.paypal.com''. In [131], 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 [132]. Such training games require algorithms to learn about users' belief in their own ability to accomplish a certain goal [133], 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

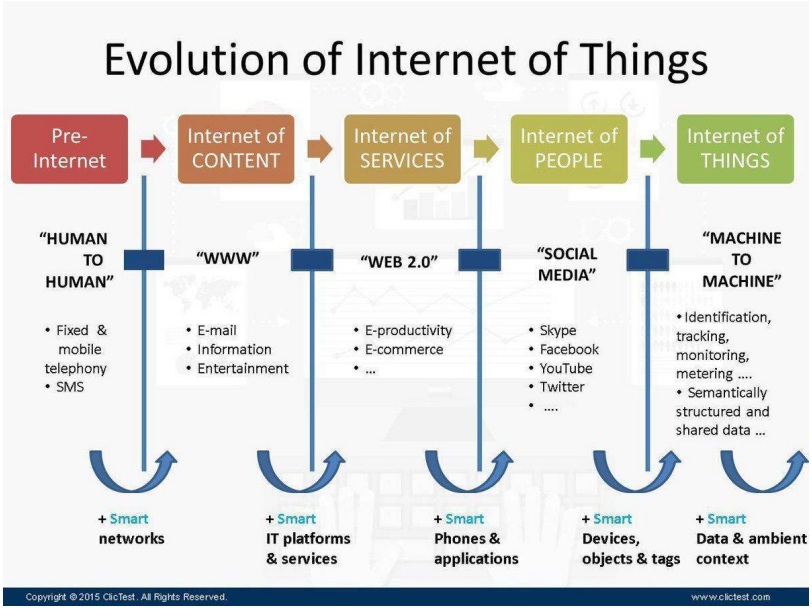

FIGURE 6. From Internet of Content to Internet of Things (Short Message Service [SMS]) [134].

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. 6 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.

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 [135], and provides real-time energy [136] and carbon footprint [137] management. Furthermore, advances in telecommunications technologies led to the development of vehicular networking applications, which enable fast data transfers between mobile devices [138]–[140].

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 [141]–[143] 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 [143]. One study [141] 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 [142].

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 [144], and artificial immune system algorithms were adopted to securely self-organize Wireless Sensor Network (WSN) ad hoc connections to serve mobile gadgets [145]. 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, in [145], 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 behavior. They used artificial immune system algorithms to classify a device's behavior as normal/abnormal. Upon detecting

abnormal behavior, 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 [146], making it impractical for cloud operators to eavesdrop.

Secure authentication mechanisms also made use of well-known biometrics and human behavior 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 [147].

One disruptive technology that can bypass legislation to promote privacy is blockchain [148]. 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 [149]–[151] to facilitate blockchain applications. In [149], 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 fulfills the end devices' access control policies, allowing automatic resource sharing between IoT devices.

The work in [150] 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 tumor 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.

In [151], 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 in [151] 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 [152]. 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" [153], China's "Made in China 2025" [154], and western countries' "Smart City" [155], 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 [155], control systems to perform real-time analysis on remote hydroelectric power plants [156], and assess reliability and safety on railway control systems [157]. 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 [158].

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 [159], [160]. 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 [161], [162]. They are used to solve profit management problems, where selling and buying to/from the grid are subject to varying energy tariffs [163].

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 [164], 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 [19]. 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 well-being 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 [165].

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 [19]. 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 [166]. 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 [167]. 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) [168]. 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 in the area of critical infrastructure protection. In [169], 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 analyzes the information received, generates a token, and forwards all the information (i.e., user privilege, command, and token) to the destination node. The node analyzes 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 [170]. 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, in [170], 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 [171]. 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.

TABLE 2. As the Internet evolves, the role of AI in cybersecurity also increases.

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

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 well-being. Not only are AI approaches being used to solve problems rationally, but also to make machines think and act like humans.

V. FUTURE CHALLENGES AND RESEARCH OPPORTUNITIES

A. THE RACE BETWEEN DEFENSE, OFFENSE, AND HUMANITY

Recent AI research advances in cybersecurity have fueled the race between the white hat (defenders) and black hat (offenders) hackers. Attackers can employ AI to mimic human behavior 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 [172]. AI has also manipulated public opinion in Venezuela by retweeting political content [173] and has affected the US presidential election by spreading tailored news [107]. 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 cyberattackers. 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. Therefore, developing fast algorithms for the AI solutions shown in Table 2 should be an active research area. 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 [174]. 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, analyzing 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 analyze cyber traffic [175], [176]. At the high end, quantum computing will be the breakthrough technology that helps solve complex computing problems. NASA's quantum computer [177] has been able to solve complex problems in a fraction of time–it is 100 million times faster [178] 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 [179].

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 [180]. In another way, researchers have started to debate whether intelligent computers should have rights

regardless of the definition of consciousness [181]. 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.

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. In particular, 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 behaviors—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. In all likelihood, 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.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions which helped us improve the content, quality, and presentation of this paper.

REFERENCES

[1] D. Venable. 2017. *Cybersecurity in 2017: When Moore's Law Attacks*. Accessed: Jun. 5, 2019. [Online]. Available: https://www.channelpartnersonline.com/blog/cybersecurity-in-2017-when-moore-s-law-attacks/

[2] S. Morgan. (Jun. 2019). Global Cybersecurity Spending Predicted to Exceed $1 Trillion From 2017–2021. Cybercrime Magazine. Accessed: Dec. 22, 2019. [Online]. Available: https://cybersecurityventures.com/cybersecurity-market-report/

[3] Statista Research Department. (Aug. 2019). *Spending on Cybersecurity in the United States From 2010 to 2018*. Accessed: Dec. 22, 2019. [Online]. Available: https://www.statista.com/statistics/615450/cybersecurity-spending-in-the-us/

[4] Wall Street. (Aug. 2018). *How Artificial Intelligence and Machine Learning Will Impact Cyber Security*. Accessed: Jan. 5, 2020. [Online]. Available: https://wall-street.com/how-artificial-intelligence-and-machine-learning-will-impact-cyber-security/

[5] D. Yuhas. (Oct. 2017). Doctors Have Trouble Diagnosing Alzheimer's. AI Doesn't. NBC News. Accessed: Dec. 25, 2019. [Online]. Available: https://www.nbcnews.com/mach/science/doctors-have-trouble-diagnosing-alzheimer-s-ai-doesn-t-ncna815561

[6] M. McFarland. (Dec. 2017). Farmers Spot Diseased Crops Faster With Artificial Intelligence. CNN Business. Accessed: Dec. 25, 2019. [Online]. Available: https://money.cnn.com/2017/12/14/technology/corn-soybean-ai-farming/index.html

[7] C. Geib. (Jan. 2018). Nasa-Funded Research Will Let Unmanned Spacecraft Think' Using AI and Blockchain. Futurism. Accessed: Dec. 20, 2019. [Online]. Available: https://futurism.com/nasa-funds-autonomous-unmanned-spacecraft

[8] E. Winick. (Dec. 2017). Lawyer-Bots are Shaking up Jobs. MIT Technology Review. Accessed: Dec. 25, 2019. [Online]. Available: https://www.technologyreview.com/s/609556/lawyer-bots-are-shaking-up-jobs/

[9] B. Morey. (Jun. 2019). Manufacturing and AI: Promises and Pitfalls. SME. Accessed: Dec. 25, 2019. [Online]. Available: https://www.sme.org/technologies/articles/2019/june/manufacturing-and-ai-promises-and-pitfalls/

[10] S. Morrow and T. Crabtree. (Aug. 2019). The Future of Cybercrime & Security. Juniper Research. Accessed: Dec. 25, 2019. [Online]. Available: https://www.juniperresearch.com/researchstore/innovation-disruption/cybercrime-security?utm_source=juniperpr&utm_campaign=pr1_thefutureofcybercrime_technology_aug19

[11] H. Taylor. (Sep. 2018). *What are Cyber Threats: How They Affect you and What to do About Them*. Accessed: Jun. 5, 2019. [Online]. Available: https://preyproject.com/blog/en/what-are-cyber-threats-how-they-affect-you-what-to-do-about-them/

[12] M. Cohen. *Zero-Day Attacks are Difficult but not Impossible to Defend Against*. Accessed: Jun. 5, 2019. [Online]. Available: https://eccitsolutions.com/zero-day-attacks-difficult-not-impossible-defend/

[13] German AV-TEST GmbH Research Institute. *Malware*. Accessed: Dec. 22, 2019. [Online]. Available: https://www.av-test.org/en/statistics/malware/

[14] W. Hall and J. Pesenti. (Oct. 2017). *Growing the Artificial Intelligence Industry in the UK*. Accessed: Dec. 30, 2019. [Online]. Available: http://ftp.shujuju.cn/platform/file/2017-10-18/782c432045784854a04e458976aef0bf.pdf

[15] C. Cath, S. Wachter, B. Mittelstadt, M. Taddeo, and L. Floridi, "Artificial intelligence and the 'good society': The US, EU, and UK approach," *Sci. Eng. Ethics*, vol. 24, no. 2, pp. 505–528, 2018.

[16] E. Brynjolfsson, D. Rock, and C. Syverson, "Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics," Nat. Bureau Econ. Res., Cambridge, MA, USA, Tech. Rep. w24001, 2017.

[17] E. Ngai, Y. Hu, Y. Wong, Y. Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," *Decis. Support Syst.*, vol. 50, no. 3, pp. 559–569, Feb. 2011.

[18] J. Borenstein. (Dec. 2018). *The Challenges of Adopting a Consistent Cybersecurity Framework in the Insurance Industry*. Accessed: Jun. 5, 2019. [Online]. Available: https://www.microsoft.com/security/blog/2018/12/20/the-challenges-of-adopting-a-consistent-cybersecurity-framework-in-the-insurance-industry/

[19] C. Alcaraz and S. Zeadally, "Critical infrastructure protection: Requirements and challenges for the 21st century," *Int. J. Crit. Infrastruct. Protection*, vol. 8, pp. 53–66, Jan. 2015.

[20] Y. Cherdantseva, P. Burnap, A. Blyth, P. Eden, K. Jones, H. Soulsby, and K. Stoddart, "A review of cyber security risk assessment methods for SCADA systems," *Comput. Secur.*, vol. 56, pp. 1–27, Feb. 2016.
[21] J. Melnick. (May 2018). *Top 10 Most Common Types of Cyber Attacks*. Accessed: Jun. 5, 2019. [Online]. Available: https://blog.netwrix.com/2018/05/15/top-10-most-common-types-of-cyber-attacks/
[22] T. S. Hyslip and T. J. Holt, "Assessing the capacity of DRDoS-for-hire services in cybercrime markets," *Deviant Behav.*, vol. 40, no. 12, pp. 1609–1625, Dec. 2019.
[23] K. Trieu and Y. Yang, "Artificial intelligence-based password brute force attacks," *Proc. MWAIS*, vol. 39, 2018. [Online]. Available: http://aisel.aisnet.org/mwais2018/39
[24] E. Owusu, J. Han, S. Das, A. Perrig, and J. Zhang, "ACCessory: Password inference using accelerometers on smartphones," in *Proc. 12th Workshop Mobile Comput. Syst. Appl. (HotMobile)*, 2012, p. 9.
[25] S. Prakashkumar, E. Murugan, R. Thiagarajan, N. Krishnaveni, and E. Babby, "Analysis of cryptography performance measures using artificial neural networking," in *Proc. Int. Conf. Emerg. Current Trends Comput. Expert Technol.* Springer, 2019, pp. 313–324.
[26] P. Brucciani. (Sep. 2018). *Why Cyber Security is So Hard*. Accessed: Jun. 5, 2019. [Online]. Available: https://medium.com/datadriveninvestor/why-cyber-security-is-so-hard-fe05921a72a0
[27] S. Roy, C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya, and Q. Wu, "A survey of game theory as applied to network security," in *Proc. 43rd Hawaii Int. Conf. Syst. Sci.*, 2010, pp. 1–10.
[28] A. Zarreh, C. Saygin, H. Wan, Y. Lee, and A. Bracho, "A game theory based cybersecurity assessment model for advanced manufacturing systems," *Procedia Manuf.*, vol. 26, pp. 1255–1264, Jan. 2018.
[29] J. Pawlick, E. Colbert, and Q. Zhu, "A game-theoretic taxonomy and survey of defensive deception for cybersecurity and privacy," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–28, Aug. 2019.
[30] W. Tushar, T. K. Saha, C. Yuen, T. Morstyn, M. D. Mcculloch, H. V. Poor, and K. L. Wood, "A motivational game-theoretic approach for peer-to-peer energy trading in the smart grid," *Appl. Energy*, vol. 243, pp. 10–20, Jun. 2019.
[31] P. Chakraborty, E. Baeyens, K. Poolla, P. P. Khargonekar, and P. Varaiya, "Sharing storage in a smart grid: A coalitional game approach," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4379–4390, Jul. 2019.
[32] R. Tang, S. Wang, and H. Li, "Game theory based interactive demand side management responding to dynamic pricing in price-based demand response of smart grids," *Appl. Energy*, vol. 250, pp. 118–130, Sep. 2019.
[33] C. T. Do, N. H. Tran, C. Hong, C. A. Kamhoua, K. A. Kwiat, E. Blasch, S. Ren, N. Pissinou, and S. S. Iyengar, "Game theory for cyber security and privacy," *ACM Comput. Surv.*, vol. 50, no. 2, p. 30, 2017.
[34] *Configuring Denial of Service Protection*. Accessed: Dec. 22, 2019. [Online]. Available: https://www.cisco.com/c/en/us/td/docs/switches/lan/catalyst6500/ios/12-2SX/configuration/guide/book/dos.pdf
[35] I. Mukhopadhyay, K. S. Gupta, D. Sen, and P. Gupta, "Heuristic intrusion detection and prevention system," in *Proc. Int. Conf. Workshop Comput. Commun. (IEMCON)*, Oct. 2015, pp. 1–7.
[36] O. Depren, M. Topallar, E. Anarim, and M. K. Ciliz, "An intelligent intrusion detection system (IDS) for anomaly and misuse detection in computer networks," *Expert Syst. Appl.*, vol. 29, no. 4, pp. 713–722, Nov. 2005.
[37] N. Duffield, P. Haffner, B. Krishnamurthy, and H. A. Ringberg, "Systems and methods for rule-based anomaly detection on IP network flow," U.S. Patent 9 258 217, Feb. 9, 2016.
[38] P. Filonov, A. Lavrentyev, and A. Vorontsov, "Multivariate industrial time series with cyber-attack simulation: Fault detection using an LSTM-based predictive data model," 2016, *arXiv:1612.06676*. [Online]. Available: https://arxiv.org/abs/1612.06676
[39] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges," *Comput. Secur.*, vol. 28, nos. 1–2, pp. 18–28, Feb. 2009.
[40] Y. S. Dai, Y. P. Xiang, and Y. Pan, "Bionic autonomic nervous systems for self-defense against DoS, spyware, malware, virus, and fishing," *ACM Trans. Auton. Adapt. Syst.*, vol. 9, no. 1, pp. 1–20, Mar. 2014.
[41] J. Dykstra and E. H. Spafford, "The case for disappearing cyber security," *Commun. ACM*, vol. 61, no. 7, pp. 40–42, Jun. 2018.
[42] N. Hampton, Z. Baig, and S. Zeadally, "Ransomware behavioural analysis on windows platforms," *J. Inf. Secur. Appl.*, vol. 40, pp. 44–51, Jun. 2018.
[43] S. Mohurle and M. Patil, "A brief study of wannacry threat: Ransomware attack 2017," *Int. J. Adv. Res. Comput. Sci.*, vol. 8, no. 5, 2017.
[44] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. London, U.K.: Pearson, 2016.
[45] A. M. Turing, "Can a machine think," *Mind*, vol. 59, no. 236, pp. 433–460, 1950.
[46] H. Gelernter, "A note on syntactic symmetry and the manipulation of formal systems by machine," *Inf. Control*, vol. 2, no. 1, pp. 80–89, Apr. 1959.
[47] A. L. Samuel, "Some studies in machine learning using the game of checkers. II—Recent progress," *IBM J. Res. Develop.*, vol. 11, no. 6, pp. 601–617, 1967.
[48] B. T. Adler, L. De Alfaro, I. Pye, and V. Raman, "Measuring author contributions to the Wikipedia," in *Proc. 4th Int. Symp. Wikis (WikiSym)*, 2008, p. 15.
[49] E. Lee, J. Woo, H. Kim, A. Mohaisen, and H. K. Kim, "You are a game bot!: Uncovering game bots in MMORPGs via self-similarity in the wild," in *Proc. Netw. Distrib. Syst. Secur. Symp.*, 2016.
[50] P. Hayati, V. Potdar, K. Chai, and A. Talevski, "Web spambot detection based on Web navigation behaviour," in *Proc. 24th IEEE Int. Conf. Adv. Inf. Netw. Appl.*, 2010, pp. 797–803.
[51] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race," in *Proc. 26th Int. Conf. World Wide Web Companion (WWW)*, 2017, pp. 963–972.
[52] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017.
[53] A. R. Kang, J. Woo, J. Park, and H. K. Kim, "Online game bot detection based on party-play log analysis," *Comput. Math. Appl.*, vol. 65, no. 9, pp. 1384–1395, May 2013.
[54] C.-F. Tsai, Y.-F. Hsu, C.-Y. Lin, and W.-Y. Lin, "Intrusion detection by machine learning: A review," *expert Syst. Appl.*, vol. 36, no. 10, pp. 11994–12000, 2009.
[55] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1153–1176, 2nd Quart., 2016.
[56] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, "A deep learning approach for network intrusion detection system," in *Proc. 9th EAI Int. Conf. Bio-Inspired Inf. Commun. Technol. (BIONETICS)*, 2016, pp. 21–26.
[57] W. L. Al-Yaseen, Z. A. Othman, and M. Z. A. Nazri, "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system," *Expert Syst. Appl.*, vol. 67, pp. 296–303, Jan. 2017.
[58] F. J. Damerau, D. E. Johnson, and M. C. Buskirk, Jr., "Automatic labeling of unlabeled text data," U.S. Patent 6 697 998, Feb. 24, 2004.
[59] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 2016.
[60] S. Zeadally and M. Tsikerdekis, "Securing Internet of Things (IoT) with machine learning," *Int. J. Commun. Syst.*, vol. 33, no. 1, p. e4169, Jan. 2020.
[61] T. Bayes, "LII. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, FRS communicated by Mr. Price, in a letter to John Canton, AMFR S," *Philos. Trans. Roy. Soc. London*, no. 53, pp. 370–418, 1763.
[62] J. Zhang, C. Chen, Y. Xiang, W. Zhou, and Y. Xiang, "Internet traffic classification by aggregating correlated Naïve Bayes predictions," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 1, pp. 5–15, Jan. 2013.
[63] S. Mukherjee and N. Sharma, "Intrusion detection using Naïve Bayes classifier with feature reduction," *Procedia Technol.*, vol. 4, pp. 119–128, Jan. 2012.
[64] S. Agrawal and J. Agrawal, "Survey on anomaly detection using data mining techniques," *Procedia Comput. Sci.*, vol. 60, pp. 708–713, Jan. 2015.
[65] K. Wang and W. Shang, "Outcome prediction of DOTA2 based on Naïve Bayes classifier," in *Proc. IEEE/ACIS 16th Int. Conf. Comput. Inf. Sci. (ICIS)*, May 2017, pp. 591–593.
[66] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, and S. Y. Philip, "Top 10 algorithms in data mining," *Knowl. Inf. Syst.*, vol. 14, no. 1, pp. 1–37, 2008.
[67] S. Sahu and B. M. Mehtre, "Network intrusion detection system using J48 Decision Tree," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Aug. 2015, pp. 2023–2026.

[68] T. P. Vuong, G. Loukas, D. Gan, and A. Bezemskij, "Decision tree-based detection of denial of service and command injection attacks on robotic vehicles," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Nov. 2015, pp. 1–6.
[69] W. W. Cohen, "Fast effective rule induction," in *Proc. Mach. Learn.*, 1995, pp. 115–123.
[70] R. Rajendran, S. V. N. Santhosh Kumar, Y. Palanichamy, and K. Arputharaj, "Detection of DoS attacks in cloud networks using intelligent rule based classification system," *Cluster Comput.*, vol. 22, no. S1, pp. 423–434, Jan. 2019.
[71] X. Chen, L. Zhang, Y. Liu, and C. Tang, "Ensemble learning methods for power system cyber-attack detection," in *Proc. IEEE 3rd Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA)*, Apr. 2018, pp. 613–616.
[72] G. Folino and F. S. Pisani, "Combining ensemble of classifiers by using genetic programming for cyber security applications," in *Proc. Eur. Conf. Appl. Evol. Comput.* Springer, 2015, pp. 54–66.
[73] E. Fix and J. L. Hodges, Jr., "Discriminatory analysis-nonparametric discrimination: Consistency properties," California Univ. Berkeley, Berkeley, CA, USA, Tech. Rep. ADA800391, 1951.
[74] M.-Y. Su, "Real-time anomaly detection systems for denial-of-service attacks by weighted k-nearest-neighbor classifiers," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3492–3498, Apr. 2011.
[75] F. Zhang, H. A. D. E. Kodituwakku, J. W. Hines, and J. Coble, "Multilayer data-driven cyber-attack detection system for industrial control systems based on network, system, and process data," *IEEE Trans. Ind. Inf.*, vol. 15, no. 7, pp. 4362–4369, Jul. 2019.
[76] J. Sakhnini, H. Karimipour, and A. Dehghantanha, "Smart grid cyber attacks detection using supervised learning and heuristic feature selection," in *Proc. IEEE 7th Int. Conf. Smart Energy Grid Eng. (SEGE)*, Aug. 2019.
[77] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
[78] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul./Aug. 2008.
[79] A. Este, F. Gringoli, and L. Salgarelli, "Support vector machines for TCP traffic classification," *Comput. Netw.*, vol. 53, no. 14, pp. 2476–2490, Sep. 2009.
[80] A. Terai, S. Abe, S. Kojima, Y. Takano, and I. Koshijima, "Cyber-attack detection for industrial control system monitoring with support vector machine based on communication profile," in *Proc. IEEE Eur. Symp. Secur. Privacy Workshops (EuroS&PW)*, Apr. 2017, pp. 132–138.
[81] K. Ghanem, F. J. Aparicio-Navarro, K. G. Kyriakopoulos, S. Lambotharan, and J. A. Chambers, "Support vector machine for network intrusion and cyber-attack detection," in *Proc. Sensor Signal Process. Defence Conf. (SSPD)*, Dec. 2017, pp. 1–5.
[82] W. S. Mcculloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biol.*, vol. 52, nos. 1–2, pp. 99–115, Jan. 1990.
[83] C. Panchev, P. Dobrev, and J. Nicholson, "Detecting port scans against mobile devices with neural networks and decision trees," in *Proc. Int. Conf. Eng. Appl. Neural Netw.* Springer, 2014, pp. 175–182.
[84] S. E. Fahlman and C. Lebiere, "The cascade-correlation learning architecture," in *Proc. Adv. Neural Inf. Process. Syst.*, 1990, pp. 524–532.
[85] A. Saied, R. E. Overill, and T. Radzik, "Detection of known and unknown DDoS attacks using Artificial Neural Networks," *Neurocomputing*, vol. 172, pp. 385–393, Jan. 2016.
[86] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biol. Cybern.*, vol. 43, no. 1, pp. 59–69, 1982.
[87] E. Corchado and Á. Herrero, "Neural visualization of network traffic data for intrusion detection," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 2042–2056, Mar. 2011.
[88] D. Atienza, A. Herrero, and E. Corchado, "Neural analysis of HTTP traffic for Web attack detection," in *Proc. Comput. Intell. Secur. Inf. Syst. Conf.* Springer, 2015, pp. 201–212.
[89] D. C. Le, A. Nur Zincir-Heywood, and M. I. Heywood, "Data analytics on network traffic flows for botnet behaviour detection," in *Proc. IEEE Symp. Ser. Comput. Intell.(SSCI)*, Dec. 2016, pp. 1–7.
[90] A. Schmidt and M. Wiegand, "A survey on hate speech detection using natural language processing," in *Proc. 5th Int. Workshop Natural Lang. Process. Social Media*, 2017, pp. 1–10.
[91] C. Van Hee, E. Lefever, B. Verhoeven, J. Mennes, B. Desmet, G. De Pauw, W. Daelemans, and V. Hoste, "Detection and fine-grained classification of cyberbullying events," in *Proc. Int. Conf. Recent Adv. Natural Language Process.*, 2015, pp. 672–680.
[92] G. G. Chowdhury, "Natural language processing," *Annu. Rev. Inf. Sci. Technol.*, vol. 37, no. 1, pp. 51–89, 2005.
[93] S. Bitam, A. Mellouk, and S. Zeadally, "Bio-inspired routing algorithms survey for vehicular ad hoc networks," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 843–867, 2nd Quart., 2014.
[94] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, 2008, pp. 160–167.
[95] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
[96] H. Shi, J. Dong, W. Wang, Y. Qian, and X. Zhang, "SSGAN: Secure steganography based on generative adversarial networks," in *Proc. Pacific Rim Conf. Multimedia*. Springer, 2017, pp. 534–544.
[97] R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning DDoS detection for consumer Internet of Things devices," in *Proc. IEEE Secur. Privacy Workshops (SPW)*, May 2018, pp. 29–35.
[98] P. Yi, T. Zhu, Q. Zhang, Y. Wu, and L. Pan, "Puppet attack: A denial of service attack in advanced metering infrastructure network," *J. Netw. Comput. Appl.*, vol. 59, pp. 325–332, Jan. 2016.
[99] H. Sedjelmaci and S. M. Senouci, "Smart grid security: A new approach to detect intruders in a smart grid neighborhood area network," in *Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM)*, 2016, pp. 6–11.
[100] Q. Niyaz, W. Sun, and A. Y. Javaid, "A deep learning based ddos detection system in software-defined networking (SDN)," 2016, *arXiv:1611.07400*. [Online]. Available: https://arxiv.org/abs/1611.07400
[101] C. Li, Y. Wu, X. Yuan, Z. Sun, W. Wang, X. Li, and L. Gong, "Detection and defense of DDoS attack–based on deep learning in openflow-based SDN," *Int. J. Commun. Syst.*, vol. 31, no. 5, p. e3497, 2018.
[102] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software defined networking," in *Proc. Int. Conf. Wireless Netw. Mobile Commun. (WINCOM)*, Oct. 2016, pp. 258–263.
[103] Q. Yan, F. R. Yu, Q. Gong, and J. Li, "Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 602–622, 1st Quart., 2015.
[104] E. Adi, Z. Baig, and P. Hingston, "Stealthy denial of service (DoS) attack modelling and detection for HTTP/2 services," *J. Netw. Comput. Appl.*, vol. 91, pp. 1–13, Aug. 2017.
[105] H. Rahmani, N. Sahli, and F. Kamoun, "Distributed denial-of-service attack detection scheme-based joint-entropy," *Secur. Commun. Netw.*, vol. 5, no. 9, pp. 1049–1061, 2012.
[106] M. Tsikerdekis and S. Zeadally, "Online deception in social media," *Commun. ACM*, vol. 57, no. 9, p. 72, 2014.
[107] C. Shao, G. L. Ciampaglia, O. Varol, A. Flammini, and F. Menczer, "The spread of fake news by social bots," pp. 96–104, 2017, *arXiv:1707.07592*. [Online]. Available: https://arxiv.org/abs/1707.07592
[108] S. Gilda, "Evaluating machine learning algorithms for fake news detection," in *Proc. IEEE 15th Student Conf. Res. Develop. (SCOReD)*, 2017, pp. 110–115.
[109] M. Spitters, P. T. Eendebak, D. T. Worm, and H. Bouma, "Threat detection in tweets with trigger patterns and contextual cues," in *Proc. IEEE Joint Intell. Secur. Inform. Conf.*, Sep. 2014, pp. 216–219.
[110] P. Dewan and P. Kumaraguru, "Towards automatic real time identification of malicious posts on Facebook," in *Proc. 13th Annu. Conf. Privacy, Secur. Trust (PST)*, Jul. 2015, pp. 85–92.
[111] Y. Liu and Y.-F. B. Wu, "Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018.
[112] N. J. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news," *Proc. Assoc. Inf. Sci. Technol.*, vol. 52, no. 1, pp. 1–4, 2015.
[113] V. Rubin, N. Conroy, Y. Chen, and S. Cornwell, "Fake news or truth? Using satirical cues to detect potentially misleading news," in *Proc. 2nd Workshop Comput. Approaches Deception Detection*, 2016, pp. 7–17.
[114] J. S. Li, L.-C. Chen, J. V. Monaco, P. Singh, and C. C. Tappert, "A comparison of classifiers and features for authorship authentication of social networking messages," *Concurrency Comput., Pract. Exper.*, vol. 29, no. 14, p. e3918, 2017.
[115] E. van der Walt and J. H. P. Eloff, "A big data science experiment—Identity deception detection," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2015, pp. 416–419.
[116] P. Badjatiya, S. Gupta, M. Gupta, and V. Varma, "Deep learning for hate speech detection in tweets," in *Proc. 26th Int. Conf. World Wide Web Companion*, 2017, pp. 759–760.

[117] Y. Erkal, M. Sezgin, and S. Gunduz, "A new cyber security alert system for Twitter," in *Proc. IEEE 14th Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2015, pp. 766–770.
[118] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on twitter," in *Proc. Collaboration, Electron. Messaging, Anti-Abuse Spam Conf. (CEAS)*, vol. 6, 2010, p. 12.
[119] M. Mowbray, "The twittering machine," in *Proc. WEBIST (2)*, 2010, pp. 299–304.
[120] J. A. Chaudhry, S. A. Chaudhry, and R. G. Rittenhouse, "Phishing attacks and defenses," *Int. J. Secur. Appl.*, vol. 10, no. 1, pp. 247–256, 2016.
[121] R. Dhamija, J. D. Tygar, and M. Hearst, "Why phishing works," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2006, pp. 581–590.
[122] M. Alsharnouby, F. Alaca, and S. Chiasson, "Why phishing still works: User strategies for combating phishing attacks," *Int. J. Hum.-Comput. Stud.*, vol. 82, pp. 69–82, 2015.
[123] M. Moghimi and A. Y. Varjani, "New rule-based phishing detection method," *Expert Syst. Appl.*, vol. 53, pp. 231–242, Jul. 2016.
[124] A. Javed, P. Burnap, and O. Rana, "Prediction of drive-by download attacks on twitter," *Inf. Process. Manage.*, vol. 56, no. 3, pp. 1133–1145, 2019.
[125] V. R. Shen, C.-S. Wei, and T. T.-Y. Juang, "Javascript malware detection using a high-level fuzzy petri net," in *Proc. Int. Conf. Mach. Learn. Cybern. (ICMLC)*, vol. 2, Jul. 2018, pp. 511–514.
[126] Z. Yi, J. Ma, L. Luo, J. Yu, and Q. Wu, "Improving javascript malware classifier's security against evasion by particle swarm optimization," in *Proc. IEEE Trustcom/BigDataSE/ISPA*, Aug. 2016, pp. 1734–1740.
[127] M. Aldwairi, M. Hasan, and Z. Balbahaith, "Detection of drive-by download attacks using machine learning approach," *Int. J. Inf. Secur. Privacy*, vol. 11, no. 4, pp. 16–28, 2017.
[128] P. Likarish, E. Jung, and I. Jo, "Obfuscated malicious javascript detection using classification techniques," in *Proc. 4th Int. Conf. Malicious Unwanted Softw. (MALWARE)*, 2009, pp. 47–54.
[129] J. Wang, Y. Xue, Y. Liu, and T. H. Tan, "JSDC: A hybrid approach for javascript malware detection and classification," in *Proc. 10th ACM Symp. Inf., Comput. Commun. Secur. (ASIA CCS)*, 2015, pp. 109–120.
[130] N. A. G. Arachchilage, S. Love, and K. Beznosov, "Phishing threat avoidance behaviour: An empirical investigation," *Comput. Hum. Behavior*, vol. 60, pp. 185–197, Jul. 2016.
[131] M. Hendrix, A. Al-Sherbaz, and V. Bloom, "Game based cyber security training: Are serious games suitable for cyber security training?" *Int. J. Serious Games*, vol. 3, no. 1, 2016.
[132] J. M. Blythe and L. Coventry, "Cyber security games: A new line of risk," in *Proc. Int. Conf. Entertainment Comput.* Springer, 2012, pp. 600–603.
[133] T. Chen, J. Hammer, and L. Dabbish, "Self-efficacy-based game design to encourage security behavior online," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, 2019, p. LBW1610.
[134] ClicTest. *Evolution of Internet of Things*. Accessed: Dec. 30, 2019. [Online]. Available: https://www.clictest.com/
[135] J. Zhu, D. S. Chan, M. S. Prabhu, P. Natarajan, H. Hu, and F. Bonomi, "Improving Web sites performance using edge servers in fog computing architecture," in *Proc. IEEE 7th Int. Symp. Service-Oriented Syst. Eng.*, Mar. 2013, pp. 320–323.
[136] M. A. Al Faruque and K. Vatanparvar, "Energy management-as-a-service over fog computing platform," *IEEE Internet Things J.*, vol. 3, no. 2, pp. 161–169, Aug. 2015.
[137] C. T. Do, N. H. Tran, C. Pham, M. G. R. Alam, J. H. Son, and C. S. Hong, "A proximal algorithm for joint resource allocation and minimizing carbon footprint in geo-distributed fog computing," in *Proc. Int. Conf. Inf. Netw. (ICOIN)*, 2015, pp. 324–329.
[138] S. Chen, J. Hu, Y. Shi, and L. Zhao, "LTE-V: A TD-LTE-based V2X solution for future vehicular network," *IEEE Internet Things J.*, vol. 3, no. 6, pp. 997–1005, Sep. 2016.
[139] Z. He, J. Cao, and X. Liu, "SDVN: Enabling rapid network innovation for heterogeneous vehicular communication," *IEEE Netw.*, vol. 30, no. 4, pp. 10–15, Jul. 2016.
[140] Y. Zhang, M. Chen, N. Guizani, D. Wu, and V. C. Leung, "SOVCAN: Safety-oriented vehicular controller area network," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 94–99, Aug. 2017.
[141] B. Kirman, C. Linehan, and S. Lawson, "Reorienting geolocation data through mischievous design," in *Funology 2*. Springer, 2018, pp. 225–240.
[142] A. Garbett, J. K. Wardman, B. Kirman, C. Linehan, and S. Lawson, "Anti-social media: Communicating risk through open data, crime maps and locative media," in *Proc. HCI Korea*, 2014, pp. 145–152.
[143] H. Chung, M. Iorga, J. Voas, and S. Lee, "Alexa, can I trust you?" *Computer*, vol. 50, no. 9, pp. 100–104, 2017.
[144] N. Kumar, R. Iqbal, S. Misra, and J. J. Rodrigues, "An intelligent approach for building a secure decentralized public key infrastructure in VANET," *J. Comput. Syst. Sci.*, vol. 81, no. 6, pp. 1042–1058, 2015.
[145] K. Saleem, N. Fisal, S. Hafizah, and R. A. Rashid, "An intelligent information security mechanism for the network layer of WSN: BIOSARP," in *Computational Intelligence in Security for Information Systems*. Springer, 2011, pp. 118–126.
[146] Y. Li, K. Gai, L. Qiu, M. Qiu, and H. Zhao, "Intelligent cryptography approach for secure distributed big data storage in cloud computing," *Inf. Sci.*, vol. 387, pp. 103–115, May 2017.
[147] A. K. Nag and D. Dasgupta, "An adaptive approach for continuous multi-factor authentication in an identity eco-system," in *Proc. 9th Annu. Cyber Inf. Secur. Res. Conf.*, 2014, pp. 65–68.
[148] S. Nakamoto. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Accessed: Jan. 1, 2020. [Online]. Available: https://bitcoin.org/bitcoin.pdf
[149] A. Outchakoucht, J. P. Leroy, and H. Es-Samaali, "Dynamic access control policy based on blockchain and machine learning for the Internet of Things," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 7, pp. 417–424, 2017.
[150] P. Mamoshina, L. Ojomoko, Y. Yanovich, A. Ostrovski, A. Botezatu, P. Prikhodko, E. Izumchenko, A. Aliper, K. Romantsov, and A. Zhebrak, "Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare," *Oncotarget*, vol. 9, no. 5, p. 5665, 2018.
[151] W. Xiong and L. Xiong, "Smart contract based data trading mode using blockchain and machine learning," *IEEE Access*, vol. 7, pp. 102331–102344, 2019.
[152] T. Sanislav, S. Zeadally, and G. D. Mois, "A cloud-integrated, multilayered, agent-based cyber-physical system architecture," *Computer*, vol. 50, no. 4, pp. 27–37, 2017.
[153] K. Henning. (Apr. 2013). *Recommendations for Implementing the Strategic Initiative Industrie 4.0*. Accessed: Jan. 1, 2020. [Online]. Available: https://www.din.de/blob/76902/e8cac883f42bf28536e7e8165993f1fd/recommendations-for-implementing-industry-4-0-data.pdf
[154] K. Briken, S. Chillas, M. Krzywdzinski, A. Marks, F. Butollo, and B. Lüthje, *Made in China 2025': Intelligent Manufacturing and Work1*, I. Grugulis, C. Smith, P. Thompson, and C. Warhust, Eds. London, U.K.: Macmillan Publishers, 2017, ch. 3, pp. 42–61, doi: 10.1057/978-1-137-61014-0_3.
[155] C. Huang and L. Liao, "An intelligent agent-based collaborative workflow for inter-enterprise PCB product design," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage.*, Dec. 2007, pp. 189–193.
[156] I. Stoian, T. Sanislav, D. Capatina, L. Miclea, H. Valean, and S. Enyedi, "Multi-agent and intelligent agents' techniques implemented in a control and supervisory telematic system," in *Proc. IEEE Int. Conf. Automat., Quality Test., Robot.*, vol. 1, May 2006, pp. 463–468.
[157] W. Nowakowski, P. Bojarczak, and Z. Łukasik, "Verification and validation of railway control systems using an expert system," in *Proc. 1st Int. Conf. Intell. Transp. Syst.* Springer, 2017, pp. 43–50.
[158] H. Wasfy, T. Wasfy, J. Peters, and R. Mahfouz, "The automation of education: How computers will take over most teaching jobs," in *Proc. ASME Int. Mech. Eng. Congr. Expo.*, 2013, p. V005T05A017.
[159] S. Wei and R. Wang, "Research on the intelligent controller for air-condition with frequency change based on fuzzy neural network," in *Proc. 2nd Int. Conf. Inf. Eng. Comput. Sci.*, Dec. 2010, pp. 1–4.
[160] H. Hagras, I. Packharn, Y. Vanderstockt, N. McNulty, A. Vadher, and F. Doctor, "An intelligent agent based approach for energy management in commercial buildings," in *Proc. IEEE Int. Conf. Fuzzy Syst. (IEEE World Congr. Comput. Intell.)*, Jun. 2008, pp. 156–162.
[161] D.-T. Huynh, D. O. Mau, and C. H. Hai, "EEO-AGA: Energy efficiency optimisation in D2D communications using adaptive genetic algorithm," in *Proc. 1st Int. Conf. Internet Things Mach. Learn. (IML)*, 2017, p. 63.
[162] Y.-H. Chen, J. Emer, and V. Sze, "Using dataflow to optimize energy efficiency of deep neural network accelerators," *IEEE Micro*, vol. 37, no. 3, pp. 12–21, 2017.
[163] A. Hajizadeh and M. A. Golkar, "Intelligent control of fuel cell distributed generation systems," in *Proc. Int. Conf. Intell. Syst. Appl. Power Syst.*, 2007, pp. 1–7.
[164] M. Roopaei, P. Rad, and K.-K. R. Choo, "Cloud of things in smart agriculture: Intelligent irrigation monitoring by thermal imaging," *IEEE Cloud Comput.*, vol. 4, no. 1, pp. 10–15, Mar. 2017.

[165] C. Alcaraz and S. Zeadally, "'Critical control system protection in the 21st century," *Computer*, vol. 46, no. 10, pp. 74–83, 2013.
[166] R. Bi, C. Zhou, and D. M. Hepburn, "Applying instantaneous SCADA data to artificial intelligence based power curve monitoring and WTG fault forecasting," in *Proc. Int. Conf. Smart Grid Clean Energy Technol. (ICSGCE)*, Oct. 2016, pp. 176–181.
[167] H. Hindy, D. Brosset, E. Bayne, A. Seeam, and X. Bellekens, "Improving SIEM for critical SCADA water infrastructures using machine learning," in *Computer Security*. Springer, 2018, pp. 3–19.
[168] L. Zhou, C. Su, Z. Li, Z. Liu, and G. P. Hancke, "Automatic fine-grained access control in SCADA by machine learning," *Future Gener. Comput. Syst.*, vol. 93, pp. 548–559, Apr. 2019.
[169] O. Rysavy, J. Rab, P. Halfar, and M. Sveda, "A formal authorization framework for networked SCADA systems," in *Proc. IEEE 19th Int. Conf. Workshops Eng. Comput.-Based Syst.*, Apr. 2012, pp. 298–302.
[170] R. Jiang, R. Lu, J. Luo, C. Lai, and X. Shen, "Efficient self-healing group key management with dynamic revocation and collusion resistance for scada in smart grid," *Secur. Commun. Netw.*, vol. 8, no. 6, pp. 1026–1039, 2015.
[171] P. L. Cavalcante, J. C. Lopez, J. F. Franco, M. J. Rider, A. V. Garcia, M. R. Malveira, L. L. Martins, and L. C. M. Direito, "Centralized self-healing scheme for electrical distribution systems," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 145–155, Aug. 2015.
[172] A. Neal and S. Kouwenhoven, and O. Sa. (Jan. 2015). Quantifying Online Advertising Fraud: Ad-Click Bots vs Humans. Oxford Bio Chronometrics. Accessed: Jan. 5, 2020. [Online]. Available: http://oxford-biochron.com/downloads/OxfordBioChron_Quantifying-Online-Advertising-Fraud_Report.pdf
[173] M. Forelle, P. Howard, A. Monroy-Hernández, and S. Savage, "Political bots and the manipulation of public opinion in Venezuela," 2015, *arXiv:1507.07109*. [Online]. Available: https://arxiv.org/abs/1507.07109
[174] M. Yousefi-Azar, L. Hamey, V. Varadharajan, and M. D. McDonnell, "Fast, automatic and scalable learning to detect Android malware," in *Proc. Int. Conf. Neural Inf. Process.* Springer, 2017, pp. 848–857.
[175] G. Xu, W. Yu, Z. Chen, H. Zhang, P. Moulema, X. Fu, and C. Lu, "A cloud computing based system for cyber security management," *Int. J. Parallel, Emergent Distrib. Syst.*, vol. 30, no. 1, pp. 29–45, 2015.
[176] G. P. Gupta and M. Kulariya, "A framework for fast and efficient cyber security network intrusion detection using apache spark," *Procedia Comput. Sci.*, vol. 93, pp. 824–831, Jan. 2016.
[177] NASA. (2018). *NASA Quantum Artificial Intelligence Laboratory (Quail)*. Accessed: Jan. 5, 2020. [Online]. Available: https://ti.arc.nasa.gov/tech/dash/groups/physics/quail/
[178] W. Herkewitz. (Dec. 2015). *Google and Nasa Say Their Quantum Computer Finally Works*. Accessed: Jan. 5, 2020. [Online]. Available: https://www.popularmechanics.com/technology/gadgets/a18475/google-nasa-d-wave-quantum-computer/
[179] Australia Computer Society. (2018). *Blockchain Challenges for Australia*. Accessed: Jun. 7, 2019. [Online]. Available: https://www.acs.org.au/insightsandpublications/reports-publications/blockchain-whitepaper.html
[180] S. Dehaene, H. Lau, and S. Kouider, "What is consciousness, and could machines have it?" *Science*, vol. 358, no. 6362, pp. 486–492, 2017.
[181] O. Carter, J. Hohwy, J. van Boxtel, V. Lamme, N. Block, C. Koch, and N. Tsuchiya, "Conscious machines: Defining questions," *Science*, vol. 359, no. 6374, p. 400, 2018.

SHERALI ZEADALLY earned his bachelor's degree in computer science from the University of Cambridge, England. He also received a doctoral degree in computer science from the University of Buckingham, England. He is currently an Associate Professor in the College of Communication and Information, University of Kentucky. His research interests include Cybersecurity, privacy, Internet of Things, computer networks, and energy-efficient networking. He is a Fellow of the British Computer Society and the Institution of Engineering Technology, England.

ERWIN ADI received the B.S. degree in computer science from The State University of New York at Stony Brook, NY, USA, in 1995, the M.S. degree in communications technology from the University of Strathclyde, Glasgow, U.K., in 1998, and the Ph.D. degree in computer and security science from Edith Cowan University, Perth, Australia, in 2017.

He is currently a Postdoctoral Fellow with the University of New South Wales Canberra at the Australian Defence Force Academy, Australia, under the research group UNSW Canberra Cyber. Prior to the current assignment, he gave lectures in network and web security at Bina Nusantara University, Jakarta, Indonesia, where his students went to the final stage of a national hacking competition. He had assumed various duties in computing solutions, integrating the digital requirements of a company's headquarter and its branch offices across the Java island. He spent five years working as a Network Engineer for several telecommunication companies in Belgium, one of which was widely recognized as one of the leading network operation centers in Europe for its dedication to promptly respond to network faults.

ZUBAIR BAIG is currently a Senior Lecturer with the School of Information Technology, Deakin University, Geelong, VIC, Australia. He has authored/coauthored over 80 journal and conference papers and book chapters. His research interests are in the areas of cyber-security, the IoT, artificial intelligence, and optimization algorithms. He is serving as an Editor for the *IET Wireless Sensor Systems Journal* and PSU - A Review-Journal. He has served on numerous technical program committees of international conferences and has delivered more than 15 keynote talks on cyber-security.

IMRAN A. KHAN was born in Rawalpindi, Punjab, Pakistan, in 1979. He received the BESE degree in telecommunications from the National University of Science and Technology, Islamabad, Punjab, Pakistan, in 2002, and the M.S. degree from the Center for Advanced Studies in Engineering, Islamabad, in 2008. He is currently pursuing the Ph.D. degree in security of the Internet of Things with Deakin University Melbourne, VIC, Australia.

His main area of research includes security of IT infrastructure, and he has a vast experience implementing various technical projects using state-of-the-art network gears and firewalls. He has been associated with the corporate industry for more than 15 years. He has operated in various senior roles in the field of telecommunications and IT security. He is currently working as a Senior Security/Network Architect in a world's leading mining organization.

Mr. Khan has achieved various technical certifications in the field of IT networking and security, including CCNA, CCNP, ITIL, VMware, Juniper security, PMP, and Cisco Certified Internet Expert (CCIE# 51926).