IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Machine Learning Threatens 5G Security

JANI SUOMALAINEN[1], ARTO JUHOLA[1], SHAHRIAR SHAHABUDDIN[2], AARNE MÄMMELÄ[3], and IJAZ AHMAD[1]

[1]VTT Technical Research Center of Finland, Espoo, Finland (e-mail: [firstname.lastname]@vtt.fi)
[2] Nokia, Oulu, Finland (e-mail: [firstname.lastname]@nokia.com)
[3]VTT Technical Research Center of Finland, Oulu, Finland (e-mail: [firstname.lastname]@vtt.fi)

Corresponding author: Jani Suomalainen (e-mail: [firstname.lastname]@vtt.fi).

"This work was supported in part by the Business Finland through the PRIORITY project."

ABSTRACT

Machine learning (ML) is expected to solve many challenges in the fifth generation (5G) of mobile networks. However, ML will also open the network to several serious cybersecurity vulnerabilities. Most of the learning in ML happens through data gathered from the environment. Un-scrutinized data will have serious consequences on machines absorbing the data to produce actionable intelligence for the network. Scrutinizing the data, on the other hand, opens privacy challenges. Unfortunately, most of the ML systems are borrowed from other disciplines that provide excellent results in small closed environments. The resulting deployment of such ML systems in 5G can inadvertently open the network to serious security challenges such as unfair use of resources, denial of service, as well as leakage of private and confidential information. Therefore, in this article we dig into the weaknesses of the most prominent ML systems that are currently vigorously researched for deployment in 5G. We further classify and survey solutions for avoiding such pitfalls of ML in 5G systems.

INDEX TERMS 5G, cybersecurity, machine learning, mobile networks, survey, threats, vulnerabilities, wireless networks.

I. INTRODUCTION

Machine learning (ML) has gained a lot of research attention in wireless networks. The main aim, similar to other research, is to improve the performance of the network or the services that use the underlying network as an enabler for the services. Furthermore, the complexity in communication networks due to increasing heterogeneity in networking equipment, end-user devices, applications, and services enforces us to automate network operations [1], [2]. Thus, automation is the main driving force behind using ML in wireless networks. However, the state of the art application of ML in wireless networks has adopted a sporadic approach where the fix-and-patch philosophy prevails. In doing so, the concepts of ML are usually borrowed from existing mature technologies such as machine vision and robotics. Such use of borrowed concepts solve the particular problem under consideration, but inadvertently create other challenges. Those challenges include the inefficient use of network resources for gathering and disseminating un-called for data, straining the processing and memory capacities of different networked nodes, and un-intentionally opening the network to security vulnerabilities.

5G will connect many aspects of society through the network ranging from critical infrastructures such as e-health, transportation, and electrical grid systems to user environments such as smart homes and handheld devices. However, there will be many security challenges within the technological enablers of 5G such as software defined networking (SDN), network functions virtualization (NFV), massive multiple-input and multiple-output (MIMO) antennas, and the diverse types of devices and services such as Internet of Things (IoT) devices and virtual reality services [3]. Since, for most of the novel services ML technologies have been sought to help minimize manual configurations or human involvement, a pertinent question arises: will ML approaches be secure or further open the network to more security vulnerabilities and challenges?

The integration of the concepts of ML and 5G can lead to potential security threats and challenges if proper consideration is not given to the underlying security concerns [18]. Some of these emerging weaknesses have been recognized by the research community and mentioned in surveys, which are listed in Table 1. However, existing surveys on 5G security have given ML threats, at most, only a brief introduction [3], [5], [10], [15] or focused on specific ML approaches [19], while surveys on ML in 5G [8], [13], [16] have handled se-

TABLE 1. Surveys in the area of 5G, machine learning, and security. Surveys are classified according to their focus areas, i.e., whether they concentrate on the security of mobile networks, applications of ML in mobile networks, security threats caused by ML, or defensive mechanisms for hardening ML. The main focus of a survey is highlighted with 'F' and briefly covered areas with '+'.

Author	Description	Scope			
		5G security	5G ML	ML threats	ML defences
Ahmad *et al.* [3] (2019)	A broad survey on security for 5G and beyond networks	F	-	+	-
Al-Garadi *et al.* [4] (2020)	Security of deep learning for internet of things (IoT)	-	-	F	F
Arfaoui *et al.* [5] (2018)	Architectural requirements for 5G security	F	+	-	-
Barreno *et al.* [6], [7] (2006, 2010)	Seminal classifications of ML threats and mitigations	-	-	F	F
Haider *et al.* [8] (2020)	Opportunities of ML for 5G	+	F	-	-
Hayat *et al.* [9] (2020)	Security for device discovery in proximity services of 5G	F	-	-	-
Khan *et al.* [10] (2019)	Survey on 5G security and privacy	F	-	-	-
Liu *et al.* [11] (2018)	A data-driven view to ML threats and mitigations	-	-	F	F
Miller *et al.* [12] (2020)	Defensive adversarial learning in deep neural networks	-	-	+	F
Morocho-Cayamcela *et al.* [13] (2019)	ML applications in 5G and beyond networks	-	F	-	-
Papernot *et al.* [14] (2018)	Systematization of ML related attacks and defenses	-	-	F	F
Sultana *et al.* [15] (2019)	ML-based intrusion detection for SDN	F	+	-	-
Zhang *et al.* [16] (2018)	Deep learning applications for wireless networks	+	F	-	-
Yuan *et al.* [17] (2019)	Classification of attacks and defences for deep learning	-	-	F	F

curity threats only briefly. On the other hand, surveys on ML security [6], [7], [12], [14], [16], [17] miss the characteristics of technologies, applications, and data flows of 5G networks. There is no survey that would combine the three dimensions – mobile networks, security, and ML – into a single balanced presentation.

In this article, we delve into the potential security challenges caused by integrating the concepts and technologies of ML into 5G, and provide possible solutions and research directions. We provide a broad survey on ML induced security threats and solutions in the scope of 5G networks. Conventionally, functionality, performance, and cost have been studied separately from dependability and security [20]. Our interdisciplinary systems approach [21], [22] combines views from the data-driven and ML oriented security research fields with the views and approaches from mobile network and 5G platform security fields. A systems approach is needed whenever a conventional reductive or analytical approach does not work since the system is not the sum of its parts because of nonlinear relationships. We show that security vulnerabilities may emerge when applying ML to 5G networks. We provide classifications of threats and solutions and identify potential attack paths and weaknesses in 5G.

This article is organised as follows: Section II provides the background on the important intersection points of ML and 5G security. The section summarizes and classifies security threats and reviews the characteristics of generic ML attacks that may affect different domains of the 5G architecture. Section III analyses security threats in 21 use cases of ML in 5G. Section IV focuses on weaknesses in 5G networks. The section explores technical and operational capabilities for detecting and exploiting inherent vulnerabilities in 5G networks and emerging technologies. Section V provides potential solutions by surveying standardization and research activities for securing data and ML processes in 5G networks. Recommendations for further work and research are discussed in Section VI. Section VII concludes the article.

II. SECURITY CHALLENGES IN ML AND 5G

5G will utilize various concepts, disciplines, and technologies of ML to not only mitigate the risks involved with human-control, but also to empower wireless networks to self-control, adapt, and heal themselves with changing user, service and traffic requirements, as well as dynamic network conditions [23]. In this vein, ML is poised to be used in almost every part of 5G networks from the physical layer to the application layer, and for diverse services using 5G networks as underlying connectivity platform, as shown in Fig. 1.

Consequently, there are huge research efforts on ML applications for 5G , as evident from the survey articles published recently [24], [25] as well as from the standardization efforts [26]. The application of ML in these areas necessitates the investigation of possible security challenges ML could pose to 5G networks. To elaborate on the arising security challenges in 5G due to the application of ML, below we briefly describe 5G security architecture, provide classification of threats, and describe generic security challenges in ML.

A. ML AND SECURITY IN 5G ARCHITECTURE

To give a high-level overview of the possible challenges, consider Fig. 1, where the applications of ML in different parts of the network are visualized. The application areas are

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031966, IEEE Access

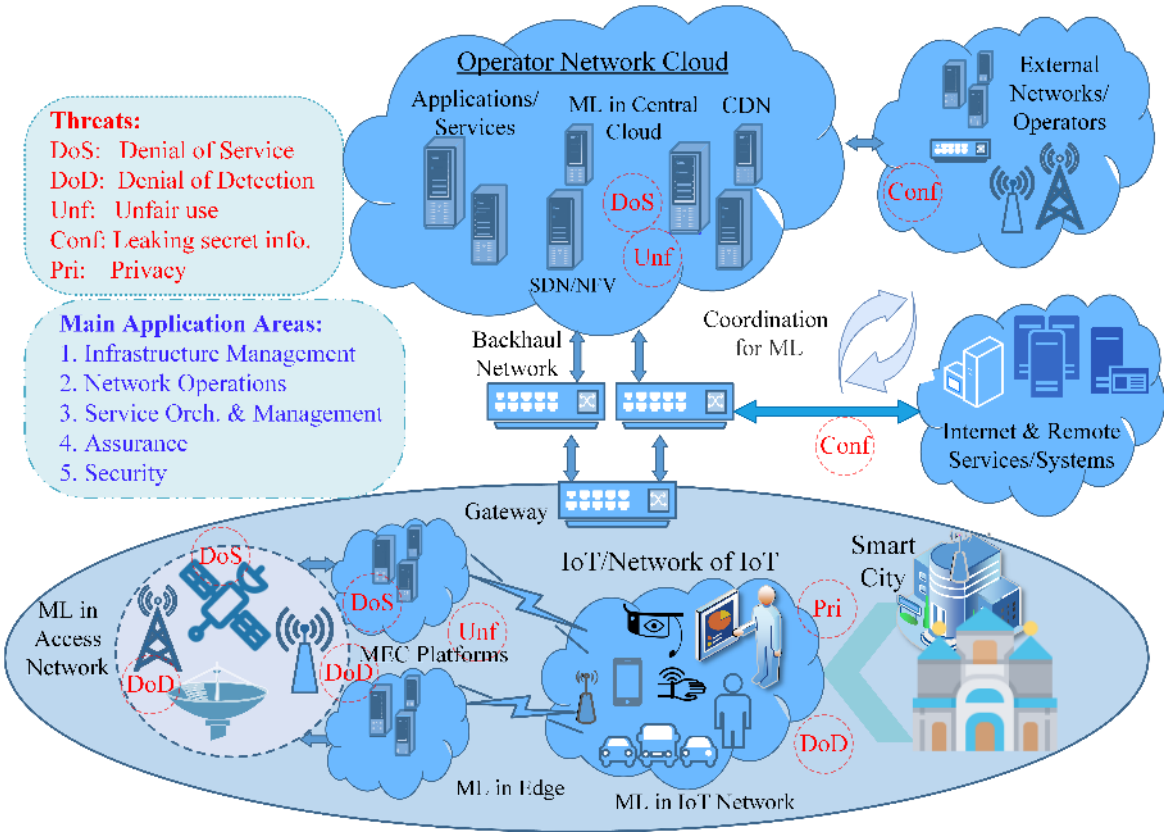

FIGURE 1. A generic network architecture of 5G using ML.

broadly categorized as 1) Infrastructure Management, 2) Network Operations, 3) Service Orchestration and Management, 4) Assurance, and 5) Security. There are many different use-cases for ML in each of these categories spanning from end-user devices and access networks to operators' central clouds. For example, ML will be applied in the access network to increase spectral efficiency or for other intelligent use of radio resources [27]; in the edge near the access network to intelligently serve latency-critical services by providing higher resources in the edge [28] and IoT [29]; in the backhaul or transport network for traffic classification [30] or improving network management with the help of SDN [31]; and for improving the performance of cloud-based services [32], [33].

The 5G security architecture has been defined in the latest 3GPP technical specification (release 15) [34] with the following main domains:

- **Network access security (I):** Comprises the set of security features that enables a user equipment (UE) to securely authenticate and access network services. Access security includes security of 3GPP and non-3GPP access technologies, and the delivery of the security context from serving network to the UE.
- **Network domain security (II):** Comprises a set of security features that enable network nodes to securely exchange signaling and user plane data.
- **User domain security (III):** Consists of security features that enable secure user access to UE.
- **Application domain security (IV):** Includes security features that enable applications (user and provider domains) to securely exchange messages.
- **Service Based Architecture (SBA) domain security (V):** Comprises of security features for network element registration, discovery, and authorization, as well as security for service-based interfaces.
- **Visibility and configurability of security (VI):** Includes features that inform users whether security features are in operation or not.

B. THREAT CATEGORIES

ML exposes various mobile assets to security threats: the configuration of infrastructure and network functions, as well as QoS levels are increasingly dependent on ML; also, information assets that are collected from 5G networks can be critical from the user privacy, operator, or customer organization confidentiality perspective.

We classify threats that ML induces into five categories. Our classification borrows, but also differs from classical cybersecurity threat classification approaches, such as spoofing, tampering, repudiation, disclosure of information, denial-of-service (DoS) and elevation of privileges (STRIDE) model [35]. Our focus is on ML vulnerabilities,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031966, IEEE Access

such as indirect tampering of models, and their 5G-specific threat manifestations. Common traditional threat categories like spoofing, tampering, repudiation, and elevation of privileges categories are seen as ways for an adversary to substantiate these threats. The main threats that ML induces include:

- **Denial-of-Service (DoS)** - causing misconfiguration, congestion or overload situations leading to unavailability of network services.
- **Denial-of-Detection (DoD)** - preventing ML from generating signals from events, attacks, or failures; enabling intrusions and other threats.
- **Unfair use or resources (Unf)** - stealing of service (e.g., routing only an adversary to an uncongested slice) or causing an extra burden or energy consumption for victims.
- **Leaking company secrets (Conf)** - adversary learns operational or business-critical information from network operators or end-user organizations. Valuable private or confidential information can reside in collected or inferred data or in the ML model itself, as noted in [36].
- **Privacy leakage (Pri)** - customer specific – e.g., user behavior revealing – sensitive parameters, data, or learned models (that can have legislative protection, such as [37]) become available to outsiders.

C. GENERIC SECURITY CHALLENGES IN ML

The basic operating principle of ML, that is, take the information of the environment (raw data) as input, process it (learning and training), and produce intelligent actionable information (classification or prediction) as output, with feedback and iterations in between, can inadvertently open security loopholes into the system. In principle, the attacks against ML are quite straightforward. An adversary can send false data to the systems which are learning or operational. An adversary may eavesdrop, intercept, or modify transmitted data. An adversary may gain access to ML processes, models, or actionable information.

1) Attacks Against ML

Attacks fooling ML can be, in general, classified using six attributes [6], [38]. Firstly, *influence* is an attribute which describes whether the attack affects training and poisons learned models, like in [39], or whether the attack tampers with learning outcomes to evade analysis, like in [40]. Secondly, *specificity* defines whether an attack is targeted and aims for mis-classifications or whether an attack is indiscriminate and affects a model's performance and reliability. Thirdly, the *security violation* attribute defines the adversary's security goal, which may be a violation of integrity, availability or privacy. Fourthly, the *frequency* attribute describes how often an attack can happen; is it a one-time event or can it happen iteratively. Fifthly, *knowledge* refers to amount of information the adversary has on the target system. In white-box attacks, the adversary knows the internals of the ML system. In black-box attacks, the adversary knows only input and outputs. Sixthly, *falsification* defines whether the goal is to get the ML model to produce false positives or false negatives.

In addition to attacks against system's operational behavior through ML processes, there are attacks that target confidential information contained in learned models or collected data. For instance, hosts learning or executing models (running in edge or in operator's cloud) may be attacked. Also, private and company-critical information can be stolen from the data masses when being transmitted or stored. Further, models can be reproduced from the training data (*inversion* attack [41]) or from the model parameters (*model extraction* attack [42]).

2) Inherent Limitations of ML Systems

The feasibility of ML depends on the quality of data. In complex and heterogeneous settings, collecting realistic and comprehensive data sets is often a challenge [43]. ML also introduces major maintenance challenges in complex settings [44]. Mixing large numbers of data sources leads to unpredictable entanglements and hidden feedback loops. Data sources may become unstable over time and have dependencies that are difficult to analyze. Similarly, models and ML-based systems may be entangled, and small changes may lead to unexpected situations and vulnerabilities. ML is by its very nature statistical, predictions are always possibilities, and in the case of many varieties of learning algorithms, the amount of error is unknown for new data. In addition, the underlying causality of unexplainable ML remains obscure, to the effect that the output *may not* reflect the intended cause, but may be something completely different with an accidental correlation with it. This kind of fault is difficult to detect since the model might still yield good results [45].

One advantage of deep-learning based ML algorithms is the capability to automatically extract features (i.e. measurable properties that are worth of being observed) from the data [46]. Unfortunately, this means that the knowledge of the individual feature's contribution to the model's predictions is lost [47]. This is a serious security handicap as well, since without this knowledge, attacker's possible additions to the training data have a better chance of remaining undetected. Insight into features is necessarily needed to spot this kind of tampering. The need for explainable AI to reveal these problems has been recognized in [45], coined as 'explain to control'.

White box adversaries – with access to models or algorithms and learning data – have the best position for attacking. However, due to *transferability* of adversarial samples [48], also adversaries with little information on victims' ML models can craft good attacks against deep learning algorithms. It is enough that the adversary can train an own substitute model for the same task and then generate adversarial samples against it.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031966, IEEE Access

TABLE 2. The use cases of machine learning in 5G domains and related threats

Use case category	Machine learning use case [26]	Domains	Threats				
			DoS	DoD	Unf	Conf	Pri
Infrastructure management	Policy-driven data centre traffic steering	I, II	x	-	x	-	-
	Handling of peak planned occurrences	II, IV	x	-	-	-	-
	Energy optimization using artificial intelligence	II, IV	x	-	x	-	-
Network operations	Policy-driven IP managed networks	II, III	x	-	-	-	x
	Radio coverage and capacity optimizations	I	x	-	x	-	-
	Intelligent software rollouts	I, II, III, V	x	x	x	-	-
	Intelligent fronthaul management and orchestration	I	x	-	x	-	-
	Elastic resource management and orchestration	II, V	x	-	x	x	-
	Application characteristic-based network operation	I, II, III	x	-	x	x	x
	AI enabled network traffic classification	II, III	x	-	x	-	x
	Automatic service and resource design framework for cloud service	II, V	x	-	x	x	-
	Intelligent time synchronization of network	II, III	x	-	x	-	-
Service orchestration and management	Context-aware voice service experience optimization	I, II, IV	x	-	-	-	x
	Intelligent network slicing management	II, IV	x	x	x	x	x
	Intelligent network carrier-managed software-defined wide area networks	II, IV	x	x	x	x	-
	Intelligent caching based on prediction of content popularity	II, IV	x	x	x	-	x
Assurance	Network fault identification and prediction	I, II	x	x	-	-	-
	Assurance of service requirements	II, IV	x	x	-	x	-
	Network fault root-cause analysis and intelligent recovery	I, II	x	x	-	x	-
Security	Policy-based network slicing for internet of things security	II	x	x	x	-	-
	Limiting profit in cyberattacks	II, III, IV, VI	-	x	x	-	-

III. ML THREATENS THE USE CASES OF 5G

ML has been proposed for various application areas of 5G. Existing work on identifying and classifying use cases include ETSI's Experiential Networked Intelligence specifications [26]. In Table 2, we map these use cases to the domains of the security architecture and to the threat categories we identified in the previous section. The use case-specific threats are then analyzed in the following subsections.

A. THREATS TO INFRASTRUCTURE MANAGEMENT

Network infrastructure, which comprises of diverse sets of networking equipment, is currently capable of having different network architectures to support diverse kinds of services. ML can support various processes for infrastructure management, including resource allocation, maintenance, and planning. For instance, ML enables intelligent load balancing [49] between data centers as well as autonomous resourcing to manage peak traffic loads and optimize energy consumption [50]. In load balancing scenarios, data is collected on link loads, and forwarding and resourcing decisions are made based on the learned optimal outcomes. High traffic load situations, such as sporting events, can be predicted using ML-algorithms, which can also assist in planning the prioritizing and resourcing.

Traffic steering and peak management can lead to DoS situations where all the traffic loads are directed towards a single target, exhausting the resources of the victim data center, and leaving other resources unused and unavailable. Policy-driven traffic steering may also lead to situations where the resources of a particular data center are consumed or purchased unfairly. Live virtual machine and virtual function migrations have also gained momentum in balancing loads in the infrastructure. At non-peak hours, ML-guided migration of virtual functions enables energy savings as idle servers can be set to a low-power state. However, an ML-based system may be spoofed to sub-optimal energy consumption, where functions remain in underutilized but high-powered servers, or to DoS situations where functions are packed into a few servers with exhausted resources.

Unmanned aerial vehicle (UAV) and satellite-based communication have been proposed [51], e.g., for access networks and backhauls in public safety scenarios. If the deployment and control of this aerial 5G infrastructure depends on ML [52], new threats could be seen towards their availability leading to risks to assets or, at worst, to personnel safety, e.g. due to crashing devices.

B. THREATS TO NETWORK OPERATIONS

ML solutions can support various management actions performed in mobile networks during run-time to enable various self-organizing capabilities [53]–[55]. ETSI use cases [26] propose, for example, a network that can manage allocations and sharing of dynamic IP addresses based on the expected needs of devices. It can also find optimal RF parameters to optimize radio coverage and capacity based on the location, load, and environmental situation of UEs. The network can also minimize disturbances by finding time periods where software updates cause minimal interference. ML can optimize the management, orchestration, and migration of fronthaul and core network functions as well as cloud services based on load estimates, predicted traffic models, utilization patterns, or an application's inspected class and characteristics. Self-organizing network capabilities supporting end-users' quality of experience could be enabled by allowing

ML to monitor thousands of KPIs and autonomously adapt networks parameters [56].

ML exposes networks, access networks, and network services for misconfigurations, which may lead to loss of availability or otherwise vulnerable states. IP allocation, software roll out, application classification and application-based network customization, as well as, time synchronization may also affect the security of user devices and expose them, for instance, to DoS attacks. Also, amounts of data, which need to be collected to optimize self-organizing networks, for example, have been identified [56] as a potential privacy issue. Furthermore, the collected data may also jeopardize an operator's secrets.

As depicted in Fig. 1, DoS attacks will be more prominent in network points where centralized decision-making happens, such as SDN control platforms, or when different users try to access the same resource, for instance in the cloud, edge or in the access network. DoD will be comparatively more threatening in access and IoT networks. IoT will be more prone to privacy leakage due to low capabilities for strong encryption, while edge and central clouds will be targets for unfair usage of resources, and clouds and communication channels will be favorite targets for the leakage of secret information. All of these challenges can be potentially exploited when ML is used without proper consideration of the security of the involved technologies on the one hand, and security weaknesses in ML techniques on the other hand.

C. THREATS TO SERVICE ORCHESTRATION

Service orchestration and management use cases automate and optimize the end-to-end network for different applications and services such as voice, IoT, or content delivery network (CDN) [57]–[60]. Networks can be customized for different applications by allocating custom resources and functions to application-specific end-to-end slices. ML can be utilized to learn and recommended optimal configuration rules, and to trigger alarms. Optimization criteria can includes attributes such as sensitivity, popularity, or cost of the content, as well as needs of users in a particular geographical location. Service orchestrations are often *multi-domain* scenarios where end-to-end service requires integration between the customer organizations' and access and core operators' (customized) networks and this makes them more dependent on each other also from security perspective.

In all the scenarios, availability is threatened due to misconfiguration or DoD. In voice and content based-services, privacy-critical end-user information is collected. In the scenarios for end-to-end and wide-area network slice management, confidential management information from the operator's network, such as health, topology, or link utilization data, may leak. Also, confidential or privacy critical information on customer behavior or assets, which have been integrated in the networks, such as operational information, procedures, as well as numbers and types of devices, may become compromised.

D. THREATS TO ASSURANCE

Assurance related use cases (considered, e.g., in [61]–[63]) analyze the network to identify and predict faults and their root-causes, as well as to allocate resources to recover from faults and to guarantee agreed service levels. Fault detection can be performed in different parts of the network, including access, backhaul and core network domains. Faults are predicted and detected by monitoring and analyzing massive amounts of data, with the help of ML. The data may originate from monitored alarms, network topology, and network service data. Root-cause analyses can then utilize e.g. decision trees to find optimal means of recovery.

There are two types of generic against assurance. First, DoD threats lead to situations where faults are not detected and, thus, corrective actions are prevented. Second, an adversary may inject false faults and alarms leading to inappropriate corrections. Both may result in DoS or other vulnerable situations. Operators can assure the behavior of their own networks and the fulfillment of service level agreements by predicting hazardous situations, detecting starvation, and by allocating and prioritizing resources for customers accordingly. However, when operators focus on service behavior in customer specific slices, there is a risk that they also infer confidential or private information on the customer.

E. THREATS TO SECURITY APPLICATIONS

5G will have many security challenges in many of its parts and technologies, as described in [3]. Since 5G systems will be much more complex compared to the previous generations, huge research efforts are dedicated to using ML for security within the novel technological concepts used in 5G, and the services that will be served by 5G networks, ranging from security of novel IoT devices and IoT services [4], [64] to virtual services in clouds [65]. In the physical layer security, ML has been demonstrated to perform well at protecting massive MIMO [66], demodulation [67], and from channel contamination in mmWave, as well as in traffic analysis and fingerprinting [68], [69]. . In the network but also in the application and UE domains, ML is applied for anomaly or signature-based intrusion detection [70] , as well as for realistic honeypot creation [71], and vulnerability scanning (e.g., scanners making use of ML for recognition of SW) [72]. Typically, ML-based security applications support expert-based or autonomous security controls. In some cases, end-users can be notified of the learned security threats with visual indicators.

The use of ML to analyze singular traffic flows and files (in an attempt to stop exploits and malware before they can wreak havoc) faces challenges which are quite similar to the ones plaguing older, signature based malware detection systems. One challenge is that the attacks, e.g., the reconnaissance and exploit delivery phases can be very stealthy. Also, the writers of new exploits/malware test their creations against the malware detectors they consider they need to thwart, ML based detectors included. Further, the providers of ML-based malware scanners need to minimize

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031966, IEEE Access

the probability of "false positives" stemming from new, legitimate software. Hence, the industry has noticed that ML models also have a "best before" date [73]. Similarly, anomaly detection-based applications require that models are continually learning or updated now and then, based on more recent training data. Unfortunately, availability and quality of relevant adversarial data for training is a challenge for ML-based attack detection applications. Realistic data is often difficult to collect as advanced attacks (that can be detected) are relatively rare. Existing public data sets, such as Network Security Laboratory-Knowledge Discovery and Data Mining (NSL-KDD) [74] and its predecessors, also have limitations [75] making them poor at detecting new attacks.

DoD is the evident security consequence of failing intrusion-detection applications. In addition, inappropriate ML-driven configurations may lead to the unfair use of resources such as quarantining innocents, or DoS situations. When ML replaces alternative security mechanisms, such as fine-grained access policing or patching of pre-5G vulnerabilities as they break backward compatibility or require extensive manual work, there is a risk that the overall security posture of the system will be eroded. In 5G, the security risks are even higher due to the integration of new technologies using ML, as elaborated in the following section.

IV. ATTACKS AND WEAKNESSES IN 5G

This section explores ML-related security attacks against 5G networks. First, we will focus on mobile network-specific vulnerabilities and attack vectors that can realize the threats, which were discussed in the previous section. Then, we will provide a deeper analysis of the security challenges and vulnerabilities, which are introduced by emerging 5G technologies. At the end of the section, we will take the opposite point of view and look at how adversaries may utilize their own ML-capabilities in attacks against traditional 5G functions.

A. NETWORK ENVIRONMENT-INDUCED WEAKNESSES

Relations between threats, attacks, and ML in the 5G mobile network context are illustrated in Fig. 2. A generic process model for autonomous systems [76] explains the role of ML. The adaptation system monitors 5G functions and executes reactions and reconfigurations after ML-supported analysis and planning. The adversary may be an outsider trying to influence monitoring data, or an insider or intruder within some monitored 5G function, or even within the control function. The adversary's goal is to change thebehavior of functions by affecting plans and execution or by evading detection during the analysis, or to cause the leakage of confidential or private data.

1) Attack Vectors in 5G

Depending on the use case, mobile networks have several potential attack surfaces, where poisoned training and evading operational data can come into play.

Network components (base stations, SDN switches, virtualized infrastructure and functions, and cloud and edge servers hosting ML functions etc.) may be intruded upon. An adversary that has successfully penetrated the first defenses (access controls, firewalls or physical security in mobile networks) can carry out different attacks [77] – man-in-the-middle, falsified data, spoofed data, etc. – when trying to influence ML-functions. As 5G cells are becoming smaller and functions are moving closer to the edge, the number of functions and data sources with less physical protection increases [78]. As data sources are administrated in different domains by different entities, their trustworthiness becomes difficult to determine by centralized ML components [79].

Open air interfaces provide a path to influence aspects such as measurements of the physical radio layer properties. Adversarial attacks against signal classifications are more powerful than classical (white noise based) jamming attacks on the wireless channel. For instance, in evasion attacks, slightly distorted signals can be misidentified by deep-learning classifiers [80]. Also, user plane integrity protection, which was introduced in 5G [34], is not mandatory feature and thus leaves the door open to tampered application layer data from UEs.

A misbehaving UE may input malicious data for ML functions which utilize information from the UE components. Vulnerabilities in network security may also enable UEs to gain access to ML functions, which do not utilize inputs directly from UEs. For instance, part of 5G and 4G access network communication is unprotected [81], [82]. This vulnerability enables man-in-the-middle attacks in 4G. Unfortunately, vulnerabilities identified in 4G remain still valid as 5G networks are backward compatible and, thus, open to downgrading or bidding down attacks [83]. For instance, unauthenticated broadcasting messages could enable rogue base stations to mislead ML-capable self-organized access networks [84]. These attacks against deep learning classifiers can be based, e.g., on malicious input, which is generated using fast gradient sign or Jacobian-based saliency map methods [85].

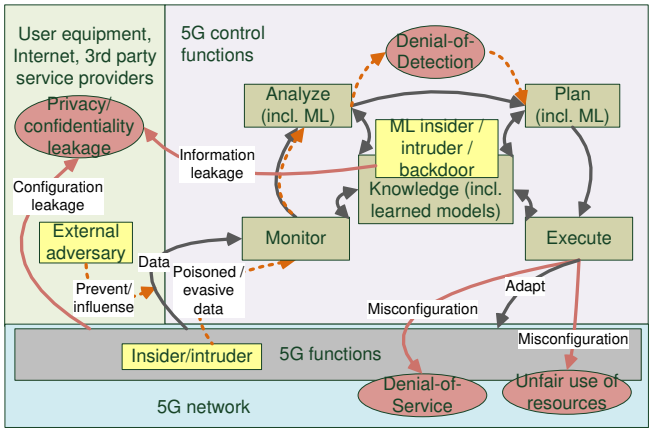

FIGURE 2. Threats in the system for autonomous 5G.

Development and supply time threats exist for ML software products, as well as devices, which are running ML and collecting data. The supplier – often a commercial or open-source third-party developer – may be careless or malicious and their systems may be intruded upon in order to implant malicious functionality or *backdoors* [86]–[88]. Backdoors are difficult to detect from ML algorithms as the malicious functionality triggers only upon input known to the adversary.

An attack on ML is likely to be a stepping stone itself, as is the case with incapacitating ML alerts. Or, the aim might be to trigger some adaptations that are beneficial for the purposes of the attacker. Often this concerns starting some reserve functionality or back-up procedure or downgrading the network to a previous generation.

2) Reconnaissance in 5G ML

In targeted white-box attacks, attackers need to understand how the ML in 5G works. A question is: how does an attacker learn about ML?

ML algorithms and software come from different suppliers. Open source components are available to all adversaries who are practicing attacks. Detailed information on commercial operator-grade solutions can be acquired only by limited parties, e.g., operators and governmental agencies, but it is a possibility that there are malicious actors within their ranks. Network management processes, including ML-related applications, reside mainly within the operator domain, allowing adversaries only indirect visibility. However, part of the data may be available through open air interfaces and some ML processes may reside in user devices and thus be available to adversaries. Adversaries inside 5G network domains may utilize network vulnerability scanning approaches and tools [77] to identify the existence of particular network services, which may then reveal information on deployed ML software, including its manufacturer and version. As ML models may be trained in other locations or by using common or synthetic data samples, an adversary may be aware also of the models which a recognized piece ofML software is executing.

External adversaries may try to resolve parameters by observing network behavior, for instance, during test attacks. However, an active testing strategy may reveal attackers methods and motives prematurely. The feedback an attacker gets from a target network is however scarce without an already gained stepping stone. An intruder aiming to tamper with ML functionality needs insider access, such as misused legitimate privileges or undetected system penetration, in order to stealthy collect, and send reconnaissance information on 5G system and ML processes.

B. SECURITY CHALLENGES IN THE LATEST TECHNOLOGIES

1) Security Challenges in Massive MIMO Systems

Massive MIMO is the most promising and disruptive technology for the 5G physical layer. In a massive MIMO system, a base station is typically equipped with a large number of antenna elements that simultaneously support a large number of users [89]. The security vulnerabilities of a massive MIMO system are divided in two categories in [90]: passive and active. In an passive attack, legitimate transmissions are eavesdropped upon. In case of an active attack, the attacker also transmits signals to disrupt or corrupt a legitimate transmission. The active attacks can be further divided into two categories: jamming attacks and pilot spoofing attacks. The goal of a jamming attack is to disrupt the transmission by sending a large amount of data towards the base station or the users. Pilot spoofing is an intelligent form of active attack where the attacker pretends to be a legitimate user by contaminating the pilots.

ML algorithms are typically used for discovering a pattern in existing data, predicting values or extracting features, which are all very useful tools to detect active adversaries. Naturally, the application of ML algorithms to secure a massive MIMO system is of great interest to the research community [91], [92]. An obvious challenge to using ML to secure a massive MIMO system is the high overhead due to the large amount of training data required by the ML algorithms. This becomes critical for large numbers of antenna streams generated in a massive MIMO system. For example, a 64-antenna base station will require separate training data for each antenna, i.e., 64 times more data and processing required than a typical intrusion detection system. Therefore, they are more vulnerable to a jamming attack when the system is already suffering from a high data overhead [92].

Unlike conventional small-scale MIMO systems, a massive MIMO system supports a large number of single antenna users. The mobility of these users can be a big challenge as the ML algorithms are typically trained for specific channel quality and characteristics. The channel characteristics can change dramatically over time and space due to the mobility of the users. A massive MIMO system trained for particular environment might not function properly in a different environment. In much of the literature, the training is done offline due to the complexity and time required for the training algorithm [93], [94]. The training algorithms, such as, backpropagation, take a considerable amount of time and hence, retraining in the field can be a challenging task.

In addition, ML schemes for massive MIMO systems also suffer from the availability of reliable data sets. Researchers often face the problem of not having broad access to real base station data while designing an algorithm. The reason is that the data sponsors are often bound by non-disclosure agreements and sharing base station data can also reveal compromising information [95]. Most researchers depend on synthesized data sets obtained by simulations and other methods. In many cases, the simulation may be based on optimistic and even unrealistic system models. For example, massive MIMO system simulations often assume perfect channel state information (CSI) availability at the transmitter or receiver, perfectly synchronized transmission and reception, uncorrelated MIMO channels etc. As a consequence, a

prevalent opinion of the research community is that experiments performed using synthetic data usually lack relevance or realism [96].

In case of supervised learning, a massive MIMO system must be trained in the absense of eavesdroppers. If the absense of an attacker cannot be guaranteed, an unsupervised learning approach needs to be adopted. Besides, it is not possible to detect a new attack when relevant training data is not available, with supervised learning methods. However, unsupervised learning algorithms are less accurate and trustworthy than supervised learning methods because the input data is not known and labeled beforehand. The user needs to interpret and label the clustered output of unsupervised learning. Due to the importance of security services, an unreliable method is not an ideal solution. In addition, the unsupervised learning algorithms are more complex compared to the supervised learning algorithms.

2) Security Challenges in SDN

SDN separates the network control functions from the data forwarding plane into a centralized control platform, serving as a central vantage point with global visibility of the network state. With programmable interfaces, the behavior of the forwarding plane can be monitored and controlled remotely, and deploying new networking functions can be simplified. However, centralizing the control plane also brings about new challenges such as security resilience and scalability. Therefore, ML approaches have been proposed to improve the resilience of the control plane under security attacks, and intelligently deploy flow forwarding rules in the data plane to avoid scalability challenges [31]. For example, in [97] decision trees [98], naive Bayes [99], and support vector networks [100] have been proposed to increase the tolerance of control platforms under different security attacks. For intelligent flow forwarding, flow feature extraction through ML has been proposed in [101], to enable application-aware policy enforcement in SDN. Similarly, ML techniques can be used to evaluation the characteristics of the flow and possible paths to minimize delays and efficiently use the available bandwidth [102].

However, an inherent limitation of SDN is overlooked in the state of the art. The SDN control platforms are still involved in fetching the data or flow features for ML algorithms to use for training and learning purposes. Since fingerprinting the SDN control platforms has been demonstrated, such systems can further make it easy for resource exhaustion attacks [103]. Similar to flow setup requests in SDN, model inversion attacks [41] in decision trees, as suggested for SDN in [97], reveal confidence values by making prediction requests to ML models. Furthermore, the most important or highly used implementations of SDN, such as OpenFlow [104], commonly used between the control and data planes, are reactive or event-driven in nature. Similarly, the north-bound APIs that have received little research attention, such as Procera [105], are also reactive. Hence, implementing proactive ML-based security measures on such reactive systems poses a significant implementation challenge. Since SDN will play an important role in different parts of 5G such as backhaul and core networks, extending security vulnerabilities in SDN with ML will directly impact the operation of the whole network.

3) Security Challenges in NFV

NFV is an 5G enabler facilitating cooperation between infrastructure providers, access and core network operators, and service providers. It eases customized deployment of network resources. In NFV, access and core network functionality are deployed as software on top of hardware infrastructure which can be shared between different tenants, network operators or application-specific slices. NFV security relies on the isolation capabilities of virtualization layer to prevent interference and information leaking between software running in different virtual machines or containers [106]. NFV management requires automated orchestration where ML has a major role. ML has been, e.g., proposed for detecting malfunctions [107], [108] and service level agreement violations [109], for classifying traffic and detecting hidden flows [110], for network QoS management [55], as well as for virtualized incident detection functionality [111], [112].

However, virtualization complicates the determination of trustworthiness and accuracy of the collected data. First, virtualization may introduce trust issues: with more cooperating parties and sharing of resource, the risk of insider attacks and information leaking increases [113]. Also, different administrative domains may not be willing to share security information on the trustworthiness of hosting platforms or hosted functions. Second, the location of monitoring probes affects both to trustworthiness and the accuracy of information [114]. The trustworthiness of data coming from migrated virtualized functions is more challenging to track as hosting servers may have different security postures at different times and location. Data coming from functions on the edge may be more likely to be compromised, due to weaker physical protection, than functions hosted on a highlysecure cloud. Probes inside virtual machines are more vulnerable as they are exposed to breaches through both the hosted functions and hosting platform. Probes outside functions, in the host or in external network hardware, are more secure but have a limited view of events.

The migration of virtual functions also has other effects on security. The models of migrated ML functions may be trained in different locations and conditions than where they are used. Consequently, unanticipated security vulnerabilities may arise. Also, confidential models may be migrated with the functions and may consequently leak if migrated to untrustworthy operation environments.

4) Security Challenges in MEC

Multi-access edge computing (MEC) facilitates 5G's low latency services by bringing computation and storage near to end-user devices. Two concepts, i.e., edge-enabled ML and ML-enabled edge are on the forefront of integrating the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031966, IEEE Access

concepts of the two technologies. Edge-enabled ML builds on the premise of low capacity of devices that force the ML processing into the resources in the edge. ML-enabled edge extends the capabilities of the edge through the methodologies of ML, in other words it makes the edge resource intelligent so it can autonomously serve the nearby end-user devices. Each of these ML-MEC integration approaches have their own security implications.

A challenge with edge-enabled ML concerns confidentiality, that is, the retaining of the confidentiality and integrity of the ML model. This is due to problems with guaranteeing isolation between the users in the edge, when this is a shared resource, and for guaranteeing the authenticity of ML functions and models migrating to the edge. No "foolproof" technology exists, and nasty surprises are always possible, as was the case with the Intel "Spectre" [115] and related "Meltdown" vulnerabilities. The ML-enabled edge functionalities are similarly vulnerable, and there may be unexpected ways they can be confused when their operation reflects external influences.

5) Security Challenges in IoT

IoT is the key area where 5G will play an important role. 5G will support IoT use cases both with radio technologies and architectural enablers [116]. Knowing that the number of devices in massive IoT can be humongous and that the data it generates will be huge, i.e., big data, using ML for efficient service provisioning will be inevitable. Massive signalling storms caused by the IoT have been recognized [117] as a challenge for backhaul and home network capacity. Similarly, large spikes of IoT triggered event information may be a challenge for the latency and capacity of ML systems. When adversary-initiated signaling spikes are rare, anomalous, and unpredictable, they may cause noise which may prevent ML from operating correctly.

Data coming from cheap, weakly protected, and often unpatched IoT types of UE are often less trusted. In some cases, small IoT devices do generate data but have no capacity for strong authentication and integrity protection. False data injection or masquerading a legitimate node in an ML-controlled system will pose serious security challenges. Network traffic analysis techniques have been successfully combined [118] with ML to profile devices and to resolve types of IoT devices. The significance of such profiling threats increases as 5G will support various special purpose devices and critical cyber-physical applications, such as healthcare and transportation.

C. ML AS AN ADVERSARIAL TOOL AGAINST 5G

Attackers may utilize their own ML solutions and products when analyzing information that is available from 5G interfaces. For instance, the wireless channel is very susceptible to eavesdropping and ML has been proposed as a means to predict transmissions and thus to find the optimal timing for jamming as, e.g., in [119]. Also, 5G control layers may leak (side-channel) information through response times [120] that may reveal knowledge about network internals or usage. Adversaries gaining access to internal 5G databases, can utilize ML techniques to infer amounts of privacy critical information in a manner that was previously unfeasible [14]. Further, ML algorithms have been shown to improve the accuracy of devices' identification via RF fingerprinting [121] and this exposes transmitting UEs for location tracking.

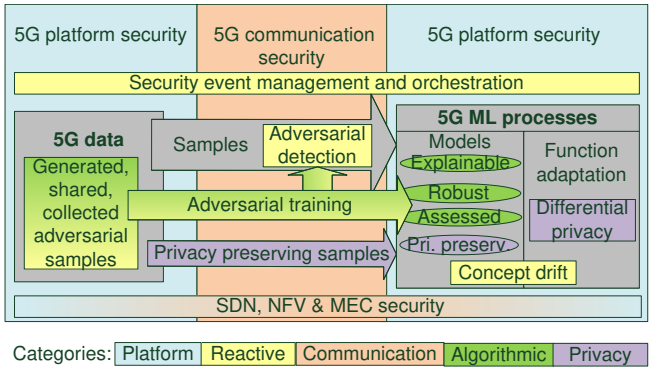

FIGURE 3. Relations between categories of solutions for securing ML systems.

V. POTENTIAL SECURITY SOLUTIONS

Security solutions to mitigate threats to ML systems can be classified to five categories as presented in Table 3. The classification combines data-driven defensive mechanism categories [11], [17], [122] and mobile network-specific security approaches. The '5G Refs' column of the table lists references to relevant standardization, applications, or research within the scope of 5G or 5G enablers. The table also identifies the domains of the security architecture (see Subsection II-A) were mechanisms are relevant and the threats (see Subsection II-B) that the mechanisms will primarily mitigate. Solutions are mapped to threats by analyzing whether they protect the confidentiality or integrity of input data, which is collected from the 5G platform (marked as 'D' in the table), whether the mechanisms protects ML algorithm, processes, or learned output (marked as 'M' in the table), or whether the mechanism is a generic approach that can protect both (marked as 'G' in the table).

Fig. 3 illustrates the central concepts of the defenses. The figure also highlights the relations between different solution categories. Platform and communication security approaches provide a protection for stored and transmitted data as well as ML processes. Emerging 5G technologies NFV, SDN, and MEC provide additional security and isolation for customer-specific ML deployments. Reactive, algorithmic, privacy enhancing technologies address specific challenges in ML systems by teaching, assessing, and supervising data sources and processes.

A. SECURITY SOLUTIONS FOR LIMITATION OF ML SYSTEMS

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031966, IEEE Access

TABLE 3. Solutions for ML induced threats

Solution Category	Mechanisms	5G Refs	Domains	Threats				
				DoS	DoD	Unf	Conf	Pri
Algorithm robustness	Adversarial training and defensive distillation	[123], [124]	I, II, IV	M	M	M	-	-
	Security assessments (penetration testing)	-	I, II, IV	M	M	M	-	-
	Explainable AI, feature engineering	[125]	I, II, IV	M	M	M	-	-
Reactive defenses	Adversarial detection over monitored samples	[126]	I, II, IV	D	D	D	-	-
	Concept drift (ML model performance monitoring)	-	I, II, IV	M	M	M	-	-
	Security information event management on 5G platforms	[127]–[132]	I, II, IV	G	G	G	G	G
	Cross-domain threat intelligence sharing	[133]	I, II, IV	G	G	G	G	G
	ML enhanced honeypots	[134]	I, II, IV	G	G	G	G	G
5G platform protection	Access control (network perimeter, operating system, physical)	[34]	I, II, IV	G	G	G	G	G
	Control and logging for ML APIs to address insider threats	-	I, II, IV, V	M	M	M	M	M
	Trusted computing for attesting integrity of network functions	[135]	I, II	G	G	G	G	G
	Slicing (SDN, NFV) to isolate application data and learning	[136]–[138]	I, II, IV	G	G	G	G	G
	Trusted security processes and personnel	-	I, II, III	G	G	G	G	G
5G communication security	Management interface security	[34]	II, V	G	G	G	G	G
	Backhaul and non-3GPP access security	[34]	I, II	D	-	-	D	D
	Access networks security - 5G-NR security	[34]	I	D	-	-	D	D
	End-to-end & application layer security	[34]	IV	D	-	-	D	D
	Continuous patching of protocol vulnerabilities	[84]	I, II	G	G	G	G	G
	Blockchains for data and inferred models	[139]	I, II, IV	G	G	-	-	-
Privacy preserving techniques	Federated learning with secure aggregation	[140]–[142]	I, III, IV	-	-	-	D	D
	Cryptography (e.g. multi-party, zero knowledge, homomorphic)	[143]	I, III, IV	-	-	-	G	G
	Differential privacy (noise to predictions and execution time)	-	I, III, IV	-	-	-	M	M
	Edge computing for locally private data	[34]	I, III, IV	-	-	-	G	G

1) Algorithm Robustness

Several techniques exist for making ML algorithms more tolerant to malicious input. In *adversarial training* [144] malicious samples are included in the training data. The approach requires that the defenders are able to collect or generate valid examples of known attacks. *Defensive distillation* trains ML models to be resilient against black box attacks [145]. In the 5G context, adversarial training has been utilized [123] to improve attack resiliency of convolutional neural network algorithms in self-organized networks. Sophisticated and targeted attacks inside 5G networks are typically quite rare. A challenge may then be how to acquire or generate realistic adversarial examples. Therefore, it is important that different domains and operators cooperate and share information on detected threats and adversarial samples. Similarly, honeypot techniques (discussed in Subsection V-B2) provide an approach to collecting adversarial samples. [146] Generative adversarial networks (essentially two 'competing' deep neural networks) can generate large amounts of adversarial training data [124].

Adversarial examples can be utilized also in *security assessments* and penetration testing. In security assessments, adversarial samples are used in testing data sets (instead of training data sets) in order to evaluate algorithms susceptibility to malicious input. Testing complex ML solutions is challenging and different tools for automation have emerged. For instance, tests and metric methodologies for evaluating how extensively the ML product has been tested have been defined [147]. Detecting backdoor attacks from ML algorithms or in general form software is challenging [148]. However, some research efforts have been made towards automating backdoor detection from deep neural networks by using activation clustering [149].

The trustworthiness of ML algorithms can be improved by minimizing number of data sources, which may be poisoned or which complicate behavior of hl the algorithms. With 'feature engineering', i.e., by analyzing the effect of features on the ML outcomes and removing features that yield no useful information, and simpler algorithms like clustering, the results vs. inputs are more explainable. As the applications of ML are becoming increasingly weighty, the explainability has received more attention [45], [125]. ML provides only probabilities of eventualities of interest. These probabilities can be improved by more mathematically rigorous approaches.

2) Reactive Defenses

Reactive solutions that monitor input or behavior (output) of ML algorithms in order to detect adversarial samples. *Concept drift* [150]–[152] is a reactive approach where the performances of ML models are continuously monitored to catch adversarial changes in behavior. The approach detects gradual changes in the accuracy of model and needs for re-calibration or re-training. *Adversarial detection* [153]–[155] approaches focus on the input of ML and utilize learning or statistical means to detect poisoned or suspicious data entries from training sets.

B. SECURITY SOLUTIONS FOR NETWORK-INDUCED WEAKNESSES

1) Platform Protection

Different technical preventive mechanisms exist for protecting and verifying the trustworthiness of equipment in 5G devices. The first layers of defenses include the physical protection in data centers and local deployments near basestations and end at the firewalls at the network boundaries. Other defensive layers include access control at the operation system, application, virtual and equipment level, as well as software configuration and trust management (including virus scanners, trusted boot, and trusted computing based attestation approaches). These solutions protect data sources and ML platforms from external attacks. Trusted computing technologies have been proposed [156] also to protect immaterial properties related to ML models.

To address insider threats, trustworthy personnel security is needed. Policies and procedures must assure that access to data and ML processes are given only for legitimate purposes. Further, access control and logging solutions are needed to guard ML interfaces. Frequent monitoring and auditing of access can then detect insiders who are misusing ML services. Personnel operating ML must be given the proper education and tools to mitigate risks of vulnerabilities due to ML misconfiguration.

2) Reactive Defenses in 5G Platforms

Security information event management solutions [127], [128], [157] utilize ML to detect ongoing security threats, anomalies, and intrusions in network elements and orchestrate automated responses. Human approval and surveillance can be part of reactive ML-based security solutions, particularly when the correct autonomous behavior in every situation cannot be trusted. Targeted monitoring enables defenses to learn more from attacks. Also, to enable monitoring systems to gain more accurate security awareness on the end-to-end situation, cross-domain data and *cyber threat intelligence sharing* [133] is needed.

An attacker needs to be wary of being tricked himself. *Honeypots* and tar-pits are an interesting playground for ML, both from the point of view of attackers and defenders. Honeypots are environments to safely learn adversarial techniques and, thus, a source for collecting realistic adversarial samples. In honeypots, ML can also have two active roles: it can help the honeypots to behave more credibly, while an attacker may use ML to tell fake and true apart. In a mobile network context, a monitoring infrastructure with high-interaction virtualized mobile device honeypots [158] has been demonstrated. The architecture for correlating honeypot information included 1) anomaly detectors for femtocell base stations, 2) malware detectors running on virtual platform on Android devices, and 3) anomaly detection in the operator command and control center (connected to the core network and analyzing control-plane data). Further, ML can also play a role in improving honeypots to be stealthier and undetectable for adversaries [71], [159]. For instance, Markov decision trees as well as reinforcement ad Q-learning have been proposed [160]–[162] for determining the optimal strategies to interact with adversaries within honeypots.

3) Communication Security

Communication security solutions are needed to protect authenticity, confidentiality, and integrity of data flowing from UEs and network functions to ML functions in operator or user organization networks. 5G provides standard security mechanisms [34] for different cases. Security in 5G new radio is based on the 5G or Extensible Authentication Protocol variant of Authentication and Key Agreement (5G-AKA or EAP-AKA') for authentication as well as the SNOW 3G, Advanced Encryption Standard - Counter Mode, or ZUC algorithms for confidentiality and integrity. Data flows between radio and core networks are protected typically with Internet Protocol Security and management interfaces towards network functions with Transmission Layer Security and Open Authentication protocols and manufacturer-specific authentication. While end-to-end learning solutions that collect data through 5G networks may utilize, e.g., the Secure Realtime Protocol or (Datagram) Transmission Layer Security to protect data.

4) Privacy Preserving Techniques

Privacy protection solutions are particularly useful when collecting application-specific information or when collecting information on UE. *Cryptographic means* – such as [143] multiparty computation, zero-knowledge argument schemes, and homomorphic encryption – can protect isolated information pieces but still enable learning models, which do not contain and, thus cannot leak, privacy-critical knowledge. Non-cryptographic solutions include *differential privacy* solutions, which protect ML results by introducing noise, e.g., to predictions or the execution time [163].

Federated learning [164] is an approach where learning is distributed to several places to improve efficiency. From the security perspective, federated learning can improve privacy, which in turn may increase the contribution and thus enable building better and more robust inference models [165]. The privacy advantage comes from secure aggregation [166]. In federated learning, raw data is not shared when collaborative models are created. The results are aggregated only when the number of data sources is sufficient and thus does not reveal privacy-sensitive information. Open challenges in federated learning are that efficiency comes at the cost of accuracy and that the cooperative parties may more easily inject backdoors into the global model because the training data is hidden [88]. Within the scope of 5G research, federated learning has been integraded [141] into the 3GPP 5G Network Data Analytics (NWDA) function. Further, blockchain based approaches [140] have been proposed to protect integrity of federated networks and to detect malicious cooperating parties. Future research challenges [142] in the 5G scope include heterogeneity of systems as well as lack of cooperation incentives and trust between different domains.

C. SECURITY SOLUTIONS FOR THE LATEST TECHNOLOGIES

1) Security Solutions for Massive MIMO

The security issues related to the usage of ML for massive MIMO systems need to be solved by not only addressing ML topics but also by modifying the system itself. The use of 64 antennas or less can provide the required spatial multiplexing gain and also reduce the data flow between the fronthaul and the baseband. Thus, the risk of a jamming attack can be reduced by reducing the antenna dimensions of a massive MIMO system.

The data overhead can also be reduced by using compression mechanisms that works with little to no loss of accuracy for a deep neural network. This is possible as there exist many redundancies with parameters of a large neural network. Two common strategies for compressing the parameters are quantization and pruning. The number of bits of the parameters can be reduced to compress the entire network. However, this requires a careful word length study to avoid performance loss. Pruning techniques reduce the number of redundant connections of a neural network.

The stability of an ML algorithm is necessary to support user mobility in massive MIMO systems. A stable ML algorithm does not deteriorate significantly when tested with a slightly different and independent dataset. A method called stability training takes perturbed samples as input to the algorithm along with the unperturbed samples and introduces a consistency constraint as an additional objective [167]. The goal of the solution is to align the outputs for unperturbed and perturbed samples. Stability training can be adopted for massive MIMO systems to support mobility. A number of inherently stable ML algorithms have been listed in [168], such as bounded support vector machine regression, regularized least square regression in a reproducing kernel Hilbert space, relative entropy regularization, maximum entropy discrimination etc., which can be explored to support mobility of massive MIMO. A naive and expensive solution is to run several ML solutions in parallel for different environments. Nevertheless, these solutions might still not be effective for very high speed moving networks, i.e., trains due to the rapid change in environments and frequent handovers.

Even though simulation-based synthesized data is frowned upon in the research community, the authors of [169] argue that radio communication presents a special case where simulation-based training data can be quite meaningful. In reality, radio signals are synthetically generated and radio channel effects are also well characterized. Hence, synthetic data generated by realistic simulation parameters and appropriate channel models can be useful for benchmarking different ML algorithms for a massive MIMO base station. We would like to note that completely relying on synthetic data is not an ideal solution either and any ML based solution should also be verified with real-world data.

An alternative solution for channel invariant active adversary detection is known as device fingerprinting. Device-dependentDevice dependent radio-metrics such as frequency and phase shift differences can be used as unique fingerprints. In [170], the authors proposed a non-parametric Bayesian method to detect and classify multiple devices in a unsupervised manner. The authors proved the effectiveness of the method against Sybil and Masquerade attacks using both simulation and experimental measurements.

2) Security Solutions for SDN

Since the main point of concern in SDN is the availability of the control plane, the concepts of ML must be used in a way not to further complicate or in the worst case , compromise availability . In general, a number of mechanisms are used to increase the availability of the control platform, such as hierarchical control plane architectures, distributing and devolving control plane functionalities, or increasing the scalability by increasing the processing capabilities and adding multiple controllers [103]. However, with ML the case will be different and simply increasing the control plane processing capabilities may not suffice [171].

One powerful capability the SDN philosophy brings to communication networks is network abstraction. A programmable network with global visibility of all resources and packets flowing through it provides an opportunity to map different services and functions according to the capability of the resources on the one hand, and monitor for security vulnerabilities and lapses on the other hand. Therefore, the deployment of ML techniques ensuring such visibility and granular control can yield the results required from ML without compromising security. To overcome the possibilities of the control plane becoming a bottleneck, an intuitive approach would be to deploy ML mechanisms after verifying the resource availability for ML processing and data exchange using the visibility of global network resource stats. For example, this could work by placing ML functions alongside specific network control functions in an edge or fog node after ensuring resources for the respective processes within those nodes, as evaluated in [172].

The selection of the mechanism of ML should be based on not only the requirements of the service or application, but also on network resources, as different ML mechanisms have different performance and network requirements in SDN [173]. For example, some mechanisms require more processing, memory, and communication rounds compared to others, and thus, in coordination with the controller will introduce more scalability and availability challenges. To deal with such problems by logically distributing the control planes, reinforcement learning mechanisms will not only facilitate coordination but also help to improve in improving resilience, as demonstrated in [129]. Reinforcement learning can also be explicitly used to improve the security of SDNs autonomously [130].

3) Security Solutions for NFV

Security and privacy requirements for ML data collection include secure communication, access controls, as well as use of ML to detect abnormally operating devices [132]. ML-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031966, IEEE Access

related NFV components in the security architecture [136], [137] include the NFV Security Controller, which orchestrates system-wide security policies, and security analytic services, which receive monitoring telemetry across NFV systems and apply ML to detect emerging threats. Security ML applications in NFV include, e.g., anomaly detection within control traffic and service chaining [131], analysis of adversarial behavior in virtual honeypots [134], as well as virtual machine and host based intrusion detection.

NFV enables customization and isolation of application-specific functions. Consequently, collected privacy and confidentiality-critical data flows can be isolated into own network slices [138]. Similarly, ML functions can be application-specific and isolated. As data flows in slices can be more homogeneous [174], input validation and algorithm robustness may also be more easily achieved. Strong isolation requires that the number of functions which are shared between slices are minimized. A challenge in slice-specific learning is that it limits ML's view of the overall situation and, hence, typically some (potentially privacy filtered) information sharing from the slices is needed.

4) Security Solutions for MEC

MEC provides inherent privacy and security protection characteristics. Data from privacy-critical applications can be stored and processed in local edge servers, which are within local administrative control and trust domains [175]. Thus critical data or applications are not necessarily exposed to threats which exists in less-trusted cloud domains. On the other hand, in situations were data backups are necessary and when MEC cannot be assured to have sufficient (e.g., physical) protection, long-term data or more critical data may be selected to be stored and processed within cloud (potentially in encrypted form). When data is shared from the edges to gain global intelligence, federated learning approaches can be utilized to minimize the amount of shared privacy critical information.

MEC can be utilized to distribute security functionality close to the end-users and access networks. To detect malicious inputs against deep learning, a distributed approach has been proposed [126] for recognizing adversarial examples. The approach decouples deep learning located in network traffic forwarding elements from the conditional generative adversarial network which is located in the mobile edge.

In general, the correct operation of ML systems depends on the availability of data. Resilient and redundant edge architectures, such as [51], can be utilized to ensure that crucial data is available for ML systems, which are located at the edge or which collect information from the edge.

VI. FUTURE RESEARCH DIRECTIONS

5G is connecting critical infrastructures through novel technological developments, where ML is poised to play an important role. However, many potential security challenges could arise, not only due to the inherent security limitations of ML, but also due to the limitations within the technologies using ML. This work discusses such challenges, their implications, and the limited possible solutions. Below some of the most pertinent future research directions which have received less research attention are discussed.

1) Security Metrics for ML in 5G

'If you cannot measure it, you cannot improve it' –a phrase often quoted from Lord Kelvin– is true also in the context of 5G cybersecurity. Operators, manufacturers, user organizations, and application providers need a comprehensive understanding of how trustworthy the network and its ML components are and how well defenses work. This awareness can be gained with formally defined metrics which measure how well available security solutions prevent identified threats. Metrics can be qualitative and follow the common criteria type of frameworks [176], or they can be quantitative as many dependability-related metrics [177]. However, there are still many gaps in research related to security metrics [178]. In the context of 5G ML, one open challenge is that affecting factors are not always observable. Spoofed ML models or even failed attempts may not be detected. Also, because the metrics are not universal and comparable, it is challenging to understand how good the system is as a whole.

We can identify and define metrics – security key performance indicators (KPIs) – for evaluating effectiveness and efficiency or trustworthiness of individual 5G ML solutions. Examples of metrics include:

- Algorithm robustness can be measured through security assessments and testing efforts as well as through the amount, freshness, and quality of adversarial samples. Samples used for training should cover known threats, including the most recent ones seen or proposed for the 5G context.
- Reactive defenses can be measured by counting the number of false positives and false negatives and true positives and negatives. The evaluation can be done against known data sets or attack libraries. Detection rate metrics are estimates based on adversarial history and do not guarantee effectiveness against new types of zero-day attacks. ETSI [179] has specified metrics for addressing the maturity of security event detection systems. Metrics can also relate to the efficiency, such as the detection time. For instance, a study of security systems for Domain Name System (DNS) security reveals that most of the systems that employ ML require hours or even days to detect threats [180].
- Platform protection and communication security mechanisms have their own strength metrics and their effectiveness can be measured by evaluating prevented or detected intrusions, events of poisoned or evasive data, or privacy breaches.

Mobile network device vendors verify and certify network equipment and function trustworthiness by using third-party test laboratories. 3GPP has developed generic and product specific-security assurance specifications and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.3031966, IEEE Access

processes [181] for evaluating security compliance of product development and product lifecycle management. The approach might be applicable for verifying data sources and ML implementations. Currently, no assurance specifications exist for ML products.

In addition to knowing the trustworthiness of an individual ML product, there is a need for solutions that can track the security situation of the whole 5G network or end-to-end service chain at the run-time. ML solutions need to track security KPIs of data sources as well as potentially adapt to changing situations and detected threats. At the network level, the situation becomes more complex as the number of data sources and ML elements increases. Future research and solutions for managing these complexities related to the trustworthiness of ML are needed.

2) Optimal Composition of Security Solutions

Inherent protection for 5G networks comes from its partially closed nature. Network components, interfaces, and functions – including ML software – are not available for everybody. 5G networks incorporate various platform and communication security solutions protecting the integrity of the platform and data and for keeping external adversaries outside. However, the size and complexity of 5G networks have left the networks partially open to advanced adversaries. Persistent adversaries will eventually find weaknesses in the large attack surface of 5G. Resource rich adversaries (including nation-level agencies, or competing operators) have the same capabilities as the defenders and may, e.g., purchase or otherwise acquire the same ML software that the defenders are using and use it for stealth testing and rehearsing attacks. Consequently, a single layer of defense is not likely to be sufficient. Vertical 5G perimeter defenses must be enforced with ML-based security applications that protect ML functions and 5G platforms from threats coming from inside, as well as with approaches for robustness and resiliency of ML algorithms.

Further research is needed to explore synergies and optimal composition of different mechanisms. It is important to determine which attacks and adversarial samples can be trusted to be filtered in perimeter defenses, and whether the remaining threats are detectable from data flows and learned models, or whether models are trainable for robustness. What compromises can be made to minimize the development and operational costs of the defenses?

One aspect to consider when doing cost-benefit evaluations, is how much resources the evolution of security solutions require. The adversaries do not rest on their laurels, and legitimate developers are likely to adopt unforeseen, progressive methods as well. For instance, when improvements on algorithm robustness were developed, new sophisticated attacks [17], [182] quickly emerged to circumvent them. Many security-related ML models have a maximum useful lifetime, and they should be updated regularly like traditional virus scanner 'signature databases'.

3) 5G Adaptations for ML Security

While solutions and research for increasing robustness and resiliency of ML algorithms exist in other domains, there is lack of research in the 5G network domain. This gap – illustrated by the lack of references in Table 3 – emphasizes the need for additional interdisciplinary and applied research. There is a need to understand how applicable different defensive solutions are with 5G's unique data flows and restrictions. Industry and the research community need to study and explore 5G adaptations to solutions such as testing [183] and certification [184], as well as explainability [185] of ML products.

4) Cross-Layer Synchronization

The research on the application of ML is mostly focused on improving a particular functionality or service in a specific layer of communication. Albeit the independence provided by the layered architecture, the use of ML in one layer can have unintentional negative consequences on another layer which can lead to security vulnerabilities. For example, intelligent spectrum sharing is gaining momentum in 5G. Hence, ML is used during the process to understand if a frequency slot is free and then to obtain it. However, different ML procedures for improving different performance metrics are used in the upper layers, such as the network or routing layer. The security of the system will require first to secure the information sharing procedure among the contending and provider peers, and second to adjust the upper layers for aspects such as secure routing, so that end-to-end security is maintained. Hence, mechanisms to synchronize different ML procedures used in different layers to avoid security lapses and to make ML solutions more robust are necessary and need further research.

5) AI-Defined Secure Networking

ML and AI have the potential to improve the trustworthiness and robustness of 5G networks. Recently, the merging of the concepts of ML in the form of ML and SDN have been proposed to bring intelligence through softwarized network functions in communication networks. However, the concepts of SDN have been limited to the OpenFlow [104] implementation of SDN. OpenFlow, no doubt, has helped implement SDN in practice, yet SDN needs more than OpenFlow offers [186]. SDN can be defined by three fundamental abstractions, i.e., forwarding, distribution and specification abstractions [186]. The forwarding abstraction should hide the underlying network complexity from applications, which OpenFlow achieves [104]. The distribution abstraction, in principle, should enable logically centralized control even though it may be physically distributed. The specification abstraction should enable applications to express a specific network behavior without delving into the network implementation. The need to identify where existing ML approaches fit into these abstractions, based on the desired improvement of a specific KPI, must be understood first. Novel ML concepts and techniques that can utilize the abstractions towards an

automated ML or AI-defined end-to-end secure network must be developed. Furthermore, ML-based security approaches that benefit from these abstractions have received very little research attention.

Interesting research questions include aspects such as whether ML could make networking threat and trust-driven. When we can automatically detect threats and attacks as well as assess the trustworthiness of ML-defined network segments, we can utilize this information in automated routing decision processes and route critical data flows through more trusted networks. More future research is needed to understand how ML can be used to infer security metrics and KPIs and to predict threat-levels in different network segments, slices, functions, or ML systems.

VII. CONCLUSION

Due to the increasing diversity in connected devices and the emergence of new services, intelligent network operations leveraging the concepts or disciplines of ML are highly researched. However, most of the state of the art takes the concepts of ML from other mature technologies such as robotics and computer vision *as it is* and use it in wireless networks such as 5G. Such a direct use of the concepts of ML in the 5G network infrastructure gives rise to many challenges, the most prominent one being compromised network security. ML opens potential vulnerabilities and attack paths against the availability and integrity of 5G services and eases user tracking and privacy violation attacks that were unfeasible with traditional adversarial methods. On the other hand, unique 5G data for learning and testing own protocols and applications in different domains, layers, and use cases necessitates solutions that are tailored for mobile networks. In this article, the challenges arising due to ML in 5G networks were discussed, followed by potential solutions to those challenges. The main objective of this work was to draw attention for future research towards secure deployment of ML techniques in 5G and future wireless networks.

ACKNOWLEDGMENT

The authors wish to thank Kimmo Halunen for the helpful discussions as well as the anonymous reviewers for their feedback.

REFERENCES

[1] M. G. Kibria, K. Nguyen, G. P. Villardi, K. Ishizu, and F. Kojima, "Next generation New Radio small cell enhancement: Architectural options, functionality and performance aspects," IEEE Wireless Communications, vol. 25, no. 4, pp. 1–9, Aug. 2018.

[2] I. Ahmad, S. Namal, M. Ylianttila, and A. Gurtov, "Towards software defined cognitive networking," in Proc. 2015 7th International Conference on New Technologies, Mobility and Security (NTMS), July 2015, pp. 1–5.

[3] I. Ahmad, S. Shahabuddin, T. Kumar, J. Okwuibe, A. Gurtov, and M. Ylianttila, "Security for 5G and beyond," IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3682–3722, Fourth Quarter 2019.

[4] M. A. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for internet of things (iot) security," IEEE Communications Surveys & Tutorials, 2020.

[5] G. Arfaoui, P. Bisson, R. Blom, R. Borgaonkar, H. Englund, E. Félix, F. Klaedtke, P. K. Nakarmi, M. Näslund, P. O'Hanlon et al., "A security architecture for 5g networks," IEEE Access, vol. 6, pp. 22 466–22 479, 2018.

[6] M. Barreno, B. Nelson, R. Sears, A. D. Joseph, and J. D. Tygar, "Can machine learning be secure?" in Proc. 2006 ACM Symposium on Information, Computer and Communications Security (ASIACS'06). New York, NY, USA: ACM, 2006, pp. 16–25.

[7] M. Barreno, B. Nelson, A. D. Joseph, and J. D. Tygar, "The security of machine learning," Machine Learning, vol. 81, no. 2, pp. 121–148, 2010.

[8] N. Haider, M. Z. Baig, and M. Imran, "Artificial intelligence and machine learning in 5g network security: Opportunities, advantages, and future research trends," arXiv preprint arXiv:2007.04490, 2020.

[9] O. Hayat, R. Ngah, Z. Kaleem, S. Z. M. Hashim, and J. J. Rodrigues, "A survey on security and privacy challenges in device discovery for next-generation systems," IEEE Access, vol. 8, pp. 84 584–84 603, 2020.

[10] R. Khan, P. Kumar, D. N. K. Jayakody, and M. Liyanage, "A survey on security and privacy of 5g technologies: Potential solutions, recent advancements, and future directions," IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp. 196–248, 2019.

[11] Q. Liu, P. Li, W. Zhao, W. Cai, S. Yu, and V. C. Leung, "A survey on security threats and defensive techniques of machine learning: A data driven view," IEEE Access, vol. 6, pp. 12 103–12 117, 2018.

[12] D. J. Miller, Z. Xiang, and G. Kesidis, "Adversarial learning targeting deep neural network classification: A comprehensive review of defenses against attacks," Proceedings of the IEEE, vol. 108, no. 3, pp. 402–433, 2020.

[13] M. E. Morocho-Cayamcela, H. Lee, and W. Lim, "Machine learning for 5G/B5G mobile and wireless communications: Potential, limitations, and future directions," IEEE Access, vol. 7, pp. 137 184–137 206, 2019.

[14] N. Papernot, P. McDaniel, A. Sinha, and M. P. Wellman, "SoK: Security and privacy in machine learning," in Proc. 2018 IEEE European Symposium on Security and Privacy (EuroS&P). IEEE, 2018, pp. 399–414.

[15] N. Sultana, N. Chilamkurti, W. Peng, and R. Alhadad, "Survey on sdn based network intrusion detection system using machine learning approaches," Peer-to-Peer Networking and Applications, vol. 12, no. 2, pp. 493–501, 2019.

[16] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," IEEE Communications Surveys & Tutorials, vol. 21, no. 3, pp. 2224–2287, Third Quarter 2019.

[17] X. Yuan, P. He, Q. Zhu, and X. Li, "Adversarial examples: Attacks and defenses for deep learning," IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 9, pp. 2805–2824, Sep. 2019.

[18] A. Qayyum, M. Usama, J. Qadir, and A. Al-Fuqaha, "Securing connected & autonomous vehicles: Challenges posed by adversarial machine learning and the way forward," IEEE Communications Surveys & Tutorials, vol. 22, no. 2, pp. 998–1026, 2020.

[19] C. Benzaid and T. Taleb, "Zsm security: Threat surface and best practices," IEEE Network, vol. 34, no. 3, pp. 124–133, 2020.

[20] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr, "Basic concepts and taxonomy of dependable and secure computing," IEEE Trans. Depend. and Secure Comput., vol. 1, no. 1, pp. 11–33, Jan.-Mar. 2004.

[21] A. F. Repko and R. Szostak, Interdisciplinary Research: Process and Theory, 4th ed. Thousand Oaks, CA: SAGE Publications, 2020.

[22] P. Checkland, Systems Thinking, Systems Practice, new ed. Chichester, UK: John Wiley & Sons, 1999.

[23] T. S. Buda, H. Assem, L. Xu, D. Raz, U. Margolin, E. Rosensweig, D. R. Lopez, M. Corici, M. Smirnov, R. Mullins, O. Uryupina, A. Mozo, B. Ordozgoiti, A. Martin, A. Alloush, P. O'Sullivan, and I. G. B. Yahia, "Can machine learning aid in delivering new use cases and scenarios in 5G?" in Proc. 2016 IEEE/IFIP Network Operations and Management Symposium (NOMS 2016), April 2016, pp. 1279–1284.

[24] Z. Xiong, Y. Zhang, D. Niyato, R. Deng, P. Wang, and L. Wang, "Deep reinforcement learning for mobile 5G and beyond: Fundamentals, applications, and challenges," IEEE Vehicular Technology Magazine, vol. 14, no. 2, pp. 44–52, June 2019.

[25] J. Wang, C. Jiang, H. Zhang, Y. Ren, K.-C. Chen, and L. Hanzo, "Thirty years of machine learning: The road to Pareto-optimal wireless networks," IEEE Communications Surveys & Tutorials, 2020.

[26] "*Experiential Networked Intelligence (ENI); ENI use cases*," European Telecommunications Standards Institute, Standard, 2019.

[27] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," IEEE Wireless Communications, vol. 24, no. 2, pp. 98–105, April 2017.

[28] J. Park, S. Samarakoon, M. Bennis, and M. Debbah, "Wireless network intelligence at the edge," Proceedings of the IEEE, vol. 107, no. 11, pp. 2204–2239, Nov. 2019.

[29] M. Mamdouh, M. A. I. Elrukhsi, and A. Khattab, "Securing the Internet of Things and wireless sensor networks via machine learning: A survey," in Proc. 2018 International Conference on Computer and Applications (ICCA), Aug 2018, pp. 215–218.

[30] Z. M. Fadlullah, F. Tang, B. Mao, N. Kato, O. Akashi, T. Inoue, and K. Mizutani, "State-of-the-art deep learning: Evolving machine intelligence toward tomorrow's intelligent network traffic control systems," IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2432–2455, Fourth Quarter 2017.

[31] J. Xie, F. R. Yu, T. Huang, R. Xie, J. Liu, C. Wang, and Y. Liu, "A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges," IEEE Communications Surveys & Tutorials, vol. 21, no. 1, pp. 393–430, First Quarter 2019.

[32] T. Pham, J. J. Durillo, and T. Fahringer, "Predicting workflow task execution time in the cloud using a two-stage machine learning approach," IEEE Transactions on Cloud Computing, vol. 8, no. 1, pp. 256–268, Jan. 2020.

[33] T. K. Rodrigues, K. Suto, H. Nishiyama, J. Liu, and N. Kato, "Machine learning meets computation and communication control in evolving edge and cloud: Challenges and future perspective," IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp. 38–67, 2019.

[34] "*Technical Specification Group Services and System Aspects (SA3); Security Architecture and Procedures for 5G System,* Release 15, TS 33.501," 3rd Generation Partnership Project (3GPP), Standard, 2018.

[35] A. Shostack, Threat Modeling: Designing for Security. Indianapolis, IN: John Wiley & Sons, 2014.

[36] N. Carlini, C. Liu, J. Kos, Ú. Erlingsson, and D. Song, "The secret sharer: Measuring unintended neural network memorization & extracting secrets," arXiv preprint arXiv:1802.08232, 2018.

[37] The European Parliament and the Council of the European Union, "Directive 2002/58/ec," 2009.

[38] O. Ibitoye, R. Abou-Khamis, A. Matrawy, and M. O. Shafiq, "The threat of adversarial attacks on machine learning in network security–a survey," arXiv preprint arXiv:1911.02621, 2019.

[39] A. N. Bhagoji, S. Chakraborty, S. Calo, and P. Mittal, "Model poisoning attacks in federated learning," in Proc. Workshop on Security in Machine Learning (SecML), collocated with 32nd Conference on Neural Information Processing Systems (NeurIPS'18), 2018.

[40] B. Biggio, I. Corona, D. Maiorca, B. Nelson, N. Šrndić, P. Laskov, G. Giacinto, and F. Roli, "Evasion attacks against machine learning at test time," in Machine Learning and Knowledge Discovery in Databases, H. Blockeel, K. Kersting, S. Nijssen, and F. Železný, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 387–402.

[41] M. Fredrikson, S. Jha, and T. Ristenpart, "Model inversion attacks that exploit confidence information and basic countermeasures," in Proc. 22nd ACM SIGSAC Conference on Computer and Communications Security. ACM, 2015, pp. 1322–1333.

[42] F. Tramèr, F. Zhang, A. Juels, M. K. Reiter, and T. Ristenpart, "Stealing machine learning models via prediction APIs," in Proc. 25th USENIX Security Symposium (USENIX Security 16), 2016, pp. 601–618.

[43] V. Gudivada, A. Apon, and J. Ding, "Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations," International Journal on Advances in Software, vol. 10, no. 1, pp. 1–20, 2017.

[44] D. Sculley, G. Holt, D. Golovin, E. Davydov, T. Phillips, D. Ebner, V. Chaudhary, and M. Young, "Machine learning: The high interest credit card of technical debt," in Proc. NIPS 2014 Workshop on Software Engineering for Machine Learning, 2014.

[45] A. Adadi and M. Berrada, "Peeking inside the black-box: A survey on explainable artificial intelligence (XAI)," IEEE Access, vol. 6, pp. 52 138–52 160, 2018.

[46] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," IEEE Communications Surveys & Tutorials, vol. 21, no. 3, pp. 2224–2287, Third Quarter 2019.

[47] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," Nature Machine Intelligence, vol. 1, no. 5, pp. 206–215, 2019.

[48] N. Papernot, P. McDaniel, and I. Goodfellow, "Transferability in machine learning: From phenomena to black-box attacks using adversarial samples," arXiv preprint arXiv:1605.07277, 2016.

[49] L.-V. Le, D. Sinh, B.-S. P. Lin, and L.-P. Tung, "Applying big data, machine learning, and SDN/NFV to 5G traffic clustering, forecasting, and management," in Proc. 2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft). IEEE, 2018, pp. 168–176.

[50] S. Zhang, N. Zhang, S. Zhou, J. Gong, Z. Niu, and X. Shen, "Energy-sustainable traffic steering for 5G mobile networks," IEEE Communications Magazine, vol. 55, no. 11, pp. 54–60, 2017.

[51] Z. Kaleem, M. Yousaf, A. Qamar, A. Ahmad, T. Q. Duong, W. Choi, and A. Jamalipour, "Uav-empowered disaster-resilient edge architecture for delay-sensitive communication," IEEE Network, vol. 33, no. 6, pp. 124–132, 2019.

[52] Q. Zhang, M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Machine learning for predictive on-demand deployment of uavs for wireless communications," in 2018 IEEE Global Communications Conference (GLOBECOM). IEEE, 2018, pp. 1–6.

[53] "*Technical Specification Group Services and System Aspects; Telecommunication management; Study on the Self-Organizing Networks (SON) for 5G networks*, release 16, TS28.861," 3rd Generation Partnership Project (3GPP), Standard, 2019.

[54] W. Jiang, M. Strufe, and H. D. Schotten, "A SON decision-making framework for intelligent management in 5G mobile networks," in Proc. 2017 3rd IEEE International Conference on Computer and Communications (ICCC). IEEE, 2017, pp. 1158–1162.

[55] I. G. B. Yahia, J. Bendriss, A. Samba, and P. Dooze, "Cognitive 5G networks: Comprehensive operator use cases with machine learning for management operations," in Proc. 2017 20th Conference on Innovations in Clouds, Internet and Networks (ICIN). IEEE, 2017, pp. 252–259.

[56] J. Moysen and L. Giupponi, "From 4G to 5G: Self-organized network management meets machine learning," Computer Communications, vol. 129, pp. 248–268, 2018.

[57] D. Laselva, M. Mattina, T. E. Kolding, J. Hui, L. Liu, and A. Weber, "Advancements of qoe assessment and optimization in mobile networks in the machine era," in Proc. 2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW). IEEE, 2018, pp. 101–106.

[58] J. Santos, T. Wauters, B. Volckaert, and F. De Turck, "Fog computing: Enabling the management and orchestration of smart city applications in 5G networks," Entropy, vol. 20, no. 1, pp. 1–20, 2018.

[59] R. Montero, F. Agraz, A. Pagès, and S. Spadaro, "End-to-end 5G service deployment and orchestration in optical networks with QoE guarantees," in Proc. 2018 20th International Conference on Transparent Optical Networks (ICTON). IEEE, 2018, pp. 1–4.

[60] E. Bastug, M. Bennis, and M. Debbah, "Living on the edge: The role of proactive caching in 5G wireless networks," IEEE Communications Magazine, vol. 52, no. 8, pp. 82–89, Aug. 2014.

[61] G. Zhu, J. Zan, Y. Yang, and X. Qi, "A supervised learning based QoS assurance architecture for 5G networks," IEEE Access, vol. 7, pp. 43 598–43 606, 2019.

[62] M. Xie, Q. Zhang, A. J. Gonzalez, P. Grønsund, P. Palacharla, and T. Ikeuchi, "Service assurance in 5G networks: A study of joint monitoring and analytics," in Proc. 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE, 2019, pp. 1–7.

[63] D. Mulvey, C. H. Foh, M. A. Imran, and R. Tafazolli, "Cell fault management using machine learning techniques," IEEE Access, vol. 7, pp. 124 514–124 539, 2019.

[64] F. Hussain, R. Hussain, S. A. Hassan, and E. Hossain, "Machine learning in IoT security: current solutions and future challenges," IEEE Communications Surveys & Tutorials, 2020.

[65] A. Sari et al., "A review of anomaly detection systems in cloud networks and survey of cloud security measures in cloud storage applications," Journal of Information Security, vol. 6, no. 02, p. 142, 2015.

[66] N. Wang, L. Jiao, and K. Zeng, "Pilot contamination attack detection for NOMA in mm-wave and massive MIMO 5G communication," in Proc. 2018 IEEE Conference on Communications and Network Security (CNS), May 2018, pp. 1–9.

[67] P. Siyari, H. Rahbari, and M. Krunz, "Lightweight machine learning for efficient frequency-offset-aware demodulation," IEEE Journal on Selected Areas in Communications, vol. 37, no. 11, pp. 2544–2558, Nov. 2019.

[68] H. D. Trinh, E. Zeydan, L. Giupponi, and P. Dini, "Detecting mobile traffic anomalies through physical control channel fingerprinting: A deep semi-supervised approach," IEEE Access, vol. 7, pp. 152 187–152 201, 2019.

[69] M. Conti, Q. Q. Li, A. Maragno, and R. Spolaor, "The dark side(-channel) of mobile devices: A survey on network traffic analysis," IEEE Communications Surveys & Tutorials, vol. 20, no. 4, pp. 2658–2713, Fourth Quarter 2018.

[70] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," IEEE Communications Surveys & Tutorials, vol. 18, no. 2, pp. 1153–1176, 2015.

[71] W. Z. A. Zakaria and M. L. M. Kiah, "A review on artificial intelligence techniques for developing intelligent honeypot," in Proc. 2012 8th International Conference on Computing Technology and Information Management (NCM and ICNIT), vol. 2. IEEE, 2012, pp. 696–701.

[72] F. Yamaguchi, F. Lindner, and K. Rieck, "Vulnerability extrapolation: Assisted discovery of vulnerabilities using machine learning," in Proc. 5th USENIX Conference on Offensive Technologies. USENIX Association, 2011, pp. 13–13.

[73] "Machine learning methods for malware detection," Kaspersky, 2020, Accessed: Feb. 11, 2020. [Online]. Available: https://media.kaspersky.com/en/enterprise-security/Kaspersky-Lab-Whitepaper-Machine-Learning.pdf

[74] "NSL-KDD dataset," Canadian Institute for Cybersecurity, Mar. 2020. [Online]. Available: https://www.unb.ca/cic/datasets/nsl.html

[75] J. McHugh, "Testing intrusion detection systems: A critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by lincoln laboratory," ACM Transactions on Information and System Security (TISSEC), vol. 3, no. 4, pp. 262–294, 2000.

[76] J. O. Kephart and D. M. Chess, "The vision of autonomic computing," IEEE Comput., vol. 36, no. 1, pp. 41–50, Jan. 2003.

[77] N. Hoque, M. H. Bhuyan, R. C. Baishya, D. K. Bhattacharyya, and J. K. Kalita, "Network attacks: Taxonomy, tools and systems," Journal of Network and Computer Applications, vol. 40, pp. 307–324, 2014.

[78] G. Chopra, R. K. Jha, and S. Jain, "Security issues in ultra dense network for 5G scenario," in Proc. 2018 10th International Conference on Communication Systems & Networks (COMSNETS). IEEE, 2018, pp. 510–512.

[79] M. Surridge, G. Correndo, K. Meacham, J. Papay, S. C. Phillips, S. Wiegand, and T. Wilkinson, "Trust modelling in 5G mobile networks," in Proc. 2018 Workshop on Security in Softwarized Networks: Prospects and Challenges, 2018, pp. 14–19.

[80] M. Sadeghi and E. G. Larsson, "Adversarial attacks on deep-learning based radio signal classification," IEEE Wireless Communications Letters, vol. 8, no. 1, pp. 213–216, Feb. 2018.

[81] S. R. Hussain, M. Echeverria, I. Karim, O. Chowdhury, and E. Bertino, "5GReasoner: A property-directed security and privacy analysis framework for 5G cellular network protocol," in Proc. 2019 ACM SIGSAC Conference on Computer and Communications Security, 2019, pp. 669–684.

[82] H. Kim, J. Lee, E. Lee, and Y. Kim, "Touching the untouchables: Dynamic security analysis of the LTE control plane," in Proc. 2019 IEEE Symposium on Security and Privacy (SP). IEEE, 2019, pp. 1153–1168.

[83] A. Shaik, R. Borgaonkar, S. Park, and J.-P. Seifert, "New vulnerabilities in 4G and 5G cellular access network protocols: Exposing device capabilities," in Proc. 12th Conference on Security and Privacy in Wireless and Mobile Networks, 2019, pp. 221–231.

[84] ——, "On the impact of rogue base stations in 4G/LTE self organising networks," in Proc. 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks. ACM, 2018, pp. 75–86.

[85] M. Usama, J. Qadir, M. A. Imran et al., "Adversarial ML attack on self organizing cellular networks," in Proc. 2019 UK/China Emerging Technologies (UCET). IEEE, 2019, pp. 1–5.

[86] T. Gu, B. Dolan-Gavitt, and S. Garg, "Badnets: Identifying vulnerabilities in the machine learning model supply chain," arXiv preprint arXiv:1708.06733, 2017.

[87] X. Chen, C. Liu, B. Li, K. Lu, and D. Song, "Targeted backdoor attacks on deep learning systems using data poisoning," arXiv preprint arXiv:1712.05526, 2017.

[88] E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, "How to backdoor federated learning," arXiv preprint arXiv:1807.00459, 2018.

[89] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," IEEE Transactions on Wireless Communications, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[90] Y. Wu, A. Khisti, C. Xiao, G. Caire, K. K. Wong, and X. Gao, "A survey of physical layer security techniques for 5G wireless networks and challenges ahead," IEEE Journal on Selected Areas in Communications, vol. 36, no. 4, pp. 679–695, April 2018.

[91] I. Ahmad, S. Shahabuddin, T. Kumar, J. Okwuibe, A. Gurtov, and M. Ylianttila, "Security for 5G and Beyond," IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3682–3722, 2019.

[92] D. Kapetanovic, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," IEEE Communications Magazine, vol. 53, no. 6, pp. 21–27, June 2015.

[93] N. Samuel, T. Diskin, and A. Wiesel, "Deep mimo detection," in 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2017, pp. 1–5.

[94] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," IEEE Transactions on Cognitive Communications and Networking, vol. 3, no. 4, pp. 563–575, 2017.

[95] S. Abt and H. Baier, "A plea for utilising synthetic data when performing machine learning based cyber-security experiments," in Proceedings of the 2014 Workshop on Artificial Intelligent and Security Workshop, 2014, pp. 37–45.

[96] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in 2010 IEEE Symposium on Security and Privacy, May 2010, pp. 305–316.

[97] S. Gangadhar and J. P. G. Sterbenz, "Machine learning aided traffic tolerance to improve resilience for software defined networks," in Proc. 9th International Workshop on Resilient Networks Design and Modeling (RNDM), Sep. 2017, pp. 1–7.

[98] J. R. Quinlan, "Induction of decision trees," Mach. Learn., vol. 1, no. 1, pp. 81–106, March 1986.

[99] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," Mach. Learn., vol. 29, no. 2, pp. 131–163, Nov. 1997.

[100] C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, Sep. 1995.

[101] Z. A. Qazi, J. Lee, T. Jin, G. Bellala, M. Arndt, and G. Noubir, "Application-awareness in SDN," SIGCOMM Comput. Commun. Rev., vol. 43, no. 4, pp. 487–488, Aug. 2013.

[102] S. T. V. Pasca, S. S. P. Kodali, and K. Kataoka, "AMPS: Application aware multipath flow routing using machine learning in SDN," in Proc. 2017 Twenty-third National Conference on Communications (NCC), March 2017, pp. 1–6.

[103] I. Ahmad, S. Namal, M. Ylianttila, and A. Gurtov, "Security in software defined networks: A survey," IEEE Communications Surveys & Tutorials, vol. 17, no. 4, pp. 2317–2346, Fourth Quarter 2015.

[104] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "Openflow: Enabling innovation in campus networks," ACM SIGCOMM Computer Communication Review, vol. 38, no. 2, pp. 69–74, April 2008.

[105] H. Kim and N. Feamster, "Improving network management with software defined networking," IEEE Communications Magazine, vol. 51, no. 2, pp. 114–119, Feb. 2013.

[106] M. D. Firoozjaei, J. P. Jeong, H. Ko, and H. Kim, "Security challenges with network functions virtualization," Future Generation Computer Systems, vol. 67, pp. 315–324, 2017.

[107] J. Ahrens, M. Strufe, L. Ahrens, and H. D. Schotten, "An AI-driven malfunction detection concept for NFV instances in 5G," arXiv preprint arXiv:1804.05796, 2018.

[108] S. Cherrared, S. Imadali, E. Fabre, and G. Gössler, "LUMEN: A global fault management framework for network virtualization environments," in Proc. 21st Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN'18). Paris, France: IEEE, Feb. 2018, pp. 1–8.

[109] J. Bendriss, I. G. B. Yahia, P. Chemouil, and D. Zeghlache, "AI for SLA management in programmable networks," in Proc. 13th International Conference on Design of Reliable Communication Networks (DRCN'17). VDE, 2017, pp. 1–8.

[110] J. Vergara-Reyes, M. C. Martinez-Ordonez, A. Ordonez, and O. M. C. Rendon, "Ip traffic classification in NFV: A benchmarking of supervised machine learning algorithms," in Proc. 2017 IEEE Colombian Conference on Communications and Computing (COLCOM), Aug 2017, pp. 1–6.

[111] G. Gardikis, K. Tzoulas, K. Tripolitis, A. Bartzas, S. Costicoglou, A. Lioy, B. Gaston, C. Fernandez, C. Davila, A. Litke et al., "Shield: A novel NFV-based cybersecurity framework," in Proc. 2017 IEEE Conference on Network Softwarization (NetSoft). IEEE, 2017, pp. 1–6.

[112] A. Sergeev, E. Ben-Sa'adon, E. Tannenbaum, and A. Saar, "Combined side-channels malware detection for NFV infrastructure," in Proc. Third Central European Cybersecurity Conference, 2019, pp. 1–2.

[113] H. Hawilo, A. Shami, M. Mirahmadi, and R. Asal, "NFV: State of the art, challenges, and implementation in next generation mobile networks (vEPC)," IEEE Network, vol. 28, no. 6, pp. 18–26, Nov.-Dec. 2014.

[114] P.-C. Lin, C.-F. Wu, and P.-H. Shih, "Optimal placement of network security monitoring functions in NFV-enabled data centers," in Proc. 2017 IEEE 7th International Symposium on Cloud and Service Computing (SC2). IEEE, 2017, pp. 9–16.

[115] P. Kocher, J. Horn, A. Fogh, D. Genkin, D. Gruss, W. Haas, M. Hamburg, M. Lipp, S. Mangard, T. Prescher, M. Schwarz, and Y. Yarom, "Spectre attacks: Exploiting speculative execution," in Proc. 2019 IEEE Symposium on Security and Privacy (SP), May 2019, pp. 1–19.

[116] S. Li, L. Da Xu, and S. Zhao, "5G Internet of Things: A survey," Journal of Industrial Information Integration, vol. 10, pp. 1–9, 2018.

[117] R. P. Jover, "Security and impact of the IoT on LTE mobile networks," in Security and Privacy in the Internet of Things (IoT): Models, Algorithms, and Implementations, F. Hu, Ed. CRC Press, 2015, vol. 6.

[118] Y. Meidan, M. Bohadana, A. Shabtai, J. D. Guarnizo, M. Ochoa, N. O. Tippenhauer, and Y. Elovici, "ProfilIoT: A machine learning approach for IoT device identification based on network traffic analysis," in Proc. Symposium on Applied Computing, 2017, pp. 506–509.

[119] T. Erpek, Y. E. Sagduyu, and Y. Shi, "Deep learning for launching and mitigating wireless jamming attacks," IEEE Transactions on Cognitive Communications and Networking, vol. 5, no. 1, pp. 2–14, March 2018.

[120] H. Cui, G. O. Karame, F. Klaedtke, and R. Bifulco, "On the fingerprinting of software-defined networks," IEEE Transactions on Information Forensics and Security, vol. 11, no. 10, pp. 2160–2173, Oct. 2016.

[121] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for RF device fingerprinting in cognitive communication networks," IEEE Journal of Selected Topics in Signal Processing, vol. 12, no. 1, pp. 160–167, 2018.

[122] L. Huang, A. D. Joseph, B. Nelson, B. I. Rubinstein, and J. D. Tygar, "Adversarial machine learning," in Proc. 4th ACM Workshop on Security and Artificial Intelligence (AISec'11). New York, NY, USA: ACM, 2011, pp. 43–58.

[123] M. Usama, J. Qadir, and A. Al-Fuqaha, "Adversarial attacks on cognitive self-organizing networks: The challenge and the way forward," in Proc. 2018 IEEE 43rd Conference on Local Computer Networks Workshops (LCN Workshops), Oct. 2018, pp. 90–97.

[124] B. Hughes, S. Bothe, H. Farooq, and A. Imran, "Generative adversarial learning for machine learning empowered self organizing 5G networks," in Proc. 2019 International Conference on Computing, Networking and Communications (ICNC). IEEE, 2019, pp. 282–286.

[125] W. Guo, "Explainable artificial intelligence (XAI) for 6G: Improving trust between human and machine," IEEE Transactions on Information Forensics and Security, vol. 11, no. 10, pp. 2160–2173, Oct. 2019.

[126] G. Li, K. Ota, M. Dong, J. Wu, and J. Li, "DeSVig: Decentralized swift vigilance against adversarial attacks in industrial artificial intelligence systems," IEEE Transactions on Industrial Informatics, vol. 16, no. 5, pp. 3267–3277, May 2020.

[127] J. Li, Z. Zhao, and R. Li, "Machine learning-based IDS for software-defined 5G network," IET Networks, vol. 7, no. 2, pp. 53–60, March 2017.

[128] I. Adam and J. Ping, "Framework for security event management in 5G," in Proc. 13th International Conference on Availability, Reliability and Security, 2018, pp. 1–7.

[129] L. S. R. Sampaio, P. H. A. Faustini, A. S. Silva, L. Z. Granville, and A. Schaeffer-Filho, "Using NFV and reinforcement learning for anomalies detection and mitigation in SDN," in Proc. 2018 IEEE Symposium on Computers and Communications (ISCC), June 2018, pp. 00 432–00 437.

[130] Y. Han, B. I. P. Rubinstein, T. Abraham, T. Alpcan, O. De Vel, S. Erfani, D. Hubczenko, C. Leckie, and P. Montague, "Reinforcement learning for autonomous defence in software-defined networking," in Decision and Game Theory for Security, L. Bushnell, R. Poovendran, and T. Başar, Eds. Cham, Switzerland: Springer International Publishing, 2018, pp. 145–165.

[131] L. Bondan, T. Wauters, B. Volckaert, F. De Turck, and L. Z. Granville, "Anomaly detection framework for SFC integrity in NFV environments," in Proc. 2017 IEEE Conference on Network Softwarization (NetSoft). IEEE, 2017, pp. 1–5.

[132] "*Experiential Networked Intelligence (ENI); ENI requirements*, etsi gr eni 002," European Telecommunications Standards Institute, Standard, 2019.

[133] X. Pan, V. Yegneswaran, Y. Chen, P. Porras, and S. Shin, "HogMap: Using SDNs to incentivize collaborative security monitoring," in Proc. 2016 ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization, 2016, pp. 7–12.

[134] A. Pastor, A. Mozo, D. R. Lopez, J. Folgueira, and A. Kapodistria, "The mouseworld, a security traffic analysis lab based on NFV/SDN," in Proc. 13th International Conference on Availability, Reliability and Security. ACM, 2018, pp. 1–6.

[135] "*Network Functions Virtualization (NFV) Release 3; NFV Security; Security and Trust Guidance*, etsi gs nfv-sec 003," European Telecommunications Standards Institute, Standard, 2014.

[136] ENISA, "Threat landscape for 5G networks," 2019.

[137] "*Network Functions Virtualization (NFV) Release 3; Security; Security Management and Monitoring specification*, ETSI GS NFV-SEC 003," European Telecommunications Standards Institute, Standard, 2017.

[138] O. Mämmelä, J. Suomalainen, K. Ahola, P. Ruuska, M. Majanen, and M. Uitto, "Micro-segmenting 5G," in Proc. IoTBDS, 2018.

[139] Y. Dai, D. Xu, S. Maharjan, Z. Chen, Q. He, and Y. Zhang, "Blockchain and deep reinforcement learning empowered intelligent 5G beyond," IEEE Network, vol. 33, no. 3, pp. 10–17, May-June 2019.

[140] Y. Liu, J. Peng, J. Kang, A. M. Iliyasu, D. Niyato, and A. A. A. El-Latif, "A secure federated learning framework for 5G networks," arXiv preprint arXiv:2005.05752, 2020.

[141] M. Isaksson and K. Norrman, "Secure federated learning in 5G mobile networks," arXiv preprint arXiv:2004.06700, 2020.

[142] Y. Liu, X. Yuan, Z. Xiong, J. Kang, X. Wang, and D. Niyato, "Federated learning for 6G communications: Challenges, methods, and future directions," arXiv preprint arXiv:2006.02931, 2020.

[143] J. Partala, L. Lovén, E. Peltonen, P. Porambage, M. Ylianttila, and T. Seppänen, "EdgeAI: A vision for privacy-preserving machine learning on the edge," in Proc. 10th Nordic Workshop on System and Network Optimization for Wireless (SNOW'19), April 2019.

[144] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," arXiv preprint arXiv:1412.6572, 2014.

[145] N. Papernot, P. McDaniel, X. Wu, S. Jha, and A. Swami, "Distillation as a defense to adversarial perturbations against deep neural networks," in Proc. 2016 IEEE Symposium on Security and Privacy (SP). IEEE, 2016, pp. 582–597.

[146] F. Younis and A. Miri, "Using honeypots in a decentralized framework to defend against adversarial machine-learning attacks," in Proc. International Conference on Applied Cryptography and Network Security. Springer, 2019, pp. 24–48.

[147] E. Breck, S. Cai, E. Nielsen, M. Salib, and D. Sculley, "The ml test score: A rubric for ml production readiness and technical debt reduction," in Proc. 2017 IEEE International Conference on Big Data (Big Data), Dec. 2017, pp. 1123–1132.

[148] C. Wysopal, C. Eng, and T. Shields, "Static detection of application backdoors," Datenschutz und Datensicherheit-DuD, vol. 34, no. 3, pp. 149–155, March 2010.

[149] B. Chen, W. Carvalho, N. Baracaldo, H. Ludwig, B. Edwards, T. Lee, I. Molloy, and B. Srivastava, "Detecting backdoor attacks on deep neural networks by activation clustering," arXiv preprint arXiv:1811.03728, 2018.

[150] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5G: How to empower SON with big data for enabling 5G," IEEE Network, vol. 28, no. 6, pp. 27–33, Nov.-Dec. 2014.

[151] G. Ditzler, M. Roveri, C. Alippi, and R. Polikar, "Learning in non-stationary environments: A survey," IEEE Computational Intelligence Magazine, vol. 10, no. 4, pp. 12–25, Nov. 2015.

[152] R. Jordaney, K. Sharad, S. K. Dash, Z. Wang, D. Papini, I. Nouretdinov, and L. Cavallaro, "Transcend: Detecting concept drift in malware classification models," in Proc. 26th USENIX Security Symposium (USENIX Security 17), 2017, pp. 625–642.

[153] J. Lu, T. Issaranon, and D. Forsyth, "Safetynet: Detecting and rejecting adversarial examples robustly," in Proc. IEEE International Conference on Computer Vision, 2017, pp. 446–454.

[154] J. H. Metzen, T. Genewein, V. Fischer, and B. Bischoff, "On detecting adversarial perturbations," arXiv preprint arXiv:1702.04267, 2017.

[155] A. N. Bhagoji, D. Cullina, C. Sitawarin, and P. Mittal, "Enhancing robustness of machine learning systems via data transformations," in Proc. 2018 52nd Annual Conference on Information Sciences and Systems (CISS). IEEE, 2018, pp. 1–5.

[156] R. Cammarota, I. Banerjee, and O. Rosenberg, "Machine learning ip protection," in Proc. 2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD), Nov. 2018, pp. 1–3.

[157] A. Aldweesh, A. Derhab, and A. Z. Emam, "Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues," Knowledge-Based Systems, vol. 189, pp. 1–19, Feb. 2020.

[158] E. Gelenbe, G. Gorbil, D. Tzovaras, S. Liebergeld, D. Garcia, M. Baltatu, and G. Lyberopoulos, "Security for smart mobile networks: The NEMESYS approach," in Proc. 2013 International Conference on Privacy and Security in Mobile Systems (PRISMS). IEEE, 2013, pp. 1–8.

[159] M. Daliran, R. Nassiri, and G. Latif-Shabgahi, "Using data analysis by deploying artificial neural networks to increase honeypot security," in Proc. 6th International Conference on Networked Computing (INC'10). IEEE, 2010, pp. 1–4.

[160] O. Hayatle, H. Otrok, and A. Youssef, "A markov decision process model for high interaction honeypots," Information Security Journal: A Global Perspective, vol. 22, no. 4, pp. 159–170, 2013.

[161] S. Dowling, M. Schukat, and E. Barrett, "Improving adaptive honeypot functionality with efficient reinforcement learning parameters for automated malware," Journal of Cyber Security Technology, vol. 2, no. 2, pp. 75–91, 2018.

[162] A. Pauna, A.-C. Iacob, and I. Bica, "QRASSH-a self-adaptive SSH honeypot driven by Q-learning," in Proc. IEEE International Conference on Communications (ICC'18). IEEE, 2018, pp. 441–446.

[163] A. D. Sarwate and K. Chaudhuri, "Signal processing and machine learning with differential privacy: Algorithms and challenges for continuous data," IEEE Signal Processing Magazine, vol. 30, no. 5, pp. 86–94, Sep. 2013.

[164] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Artificial Intelligence and Statistics. PMLR, 2017, pp. 1273–1282.

[165] W. Y. B. Lim, N. C. Luong, D. T. Hoang, Y. Jiao, Y.-C. Liang, Q. Yang, D. Niyato, and C. Miao, "Federated learning in mobile edge networks: A comprehensive survey," IEEE Communications Surveys & Tutorials, 2020.

[166] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical secure aggregation for privacy-preserving machine learning," in Proc. 2017 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2017, pp. 1175–1191.

[167] S. Zheng, Y. Song, T. Leung, and I. Goodfellow, "Improving the robustness of deep neural networks via stability training," in Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 4480–4488.

[168] H. Xu, C. Caramanis, and S. Mannor, "Sparse algorithms are not stable: A no-free-lunch theorem," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 1, pp. 187–193, Jan. 2011.

[169] T. J. O'Shea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in International conference on engineering applications of neural networks. Springer, 2016, pp. 213–226.

[170] N. T. Nguyen, G. Zheng, Z. Han, and R. Zheng, "Device fingerprinting to enhance wireless security using nonparametric bayesian method," in Proc. 2011 IEEE INFOCOM, April 2011, pp. 1404–1412.

[171] Guang Yao, Jun Bi, and Luyi Guo, "On the cascading failures of multi-controllers in software defined networks," in Proc. 21st IEEE International Conference on Network Protocols (ICNP'13), Oct 2013, pp. 1–2.

[172] Q. Shafi, A. Basit, S. Qaisar, A. Koay, and I. Welch, "Fog-assisted SDN controlled framework for enduring anomaly detection in an IoT network," IEEE Access, vol. 6, pp. 73 713–73 723, 2018.

[173] Y. Zhao, Y. Li, X. Zhang, G. Geng, W. Zhang, and Y. Sun, "A survey of networking applications applying the software defined networking concept based on machine learning," IEEE Access, vol. 7, pp. 95 397–95 417, 2019.

[174] J. Suomalainen, K. Ahola, M. Majanen, O. Mämmelä, and P. Ruuska, "Security awareness in software-defined multi-domain 5G networks," Future Internet, vol. 10, no. 3, pp. 1–24, 2018.

[175] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, fog et al.: A survey and analysis of security threats and challenges," Future Generation Computer Systems, vol. 78, pp. 680–698, Jan. 2018.

[176] J. Tierney and T. Boswell, "Common criteria: Origins and overview," in Smart Cards, Tokens, Security and Applications, K. Mayes and K. Markantonakis, Eds. Springer, 2017, pp. 193–216.

[177] M. Al-Kuwaiti, N. Kyriakopoulos, and S. Hussein, "A comparative analysis of network dependability, fault-tolerance, reliability, security, and survivability," IEEE Communications Surveys & Tutorials, vol. 11, no. 2, pp. 106–124, Second Quarter 2009.

[178] R. Diesch, M. Pfaff, and H. Krcmar, "Prerequisite to measure information security," in Proc. 4th International Conference on Information Systems Security and Privacy. SciTePress, 2018, pp. 207–201.

[179] "*Information Security Indicators (ISI); Key Performance Security Indicators (KPSI) to evaluate the maturity of security event detection*, ETSI GS ISI 003," European Telecommunications Standards Institute, Standard, Nov. 2018.

[180] S. Torabi, A. Boukhtouta, C. Assi, and M. Debbabi, "Detecting Internet abuse by analyzing passive DNS traffic: A survey of implemented systems," IEEE Communications Surveys & Tutorials, vol. 20, no. 4, pp. 3389–3415, Fourth Quarter 2018.

[181] "Security Assurance Methodology (SCAS) for 3GPP network products. TR 33.916." 3rd Generation Partnership Project (3GPP), Standard, 2019.

[182] A. Athalye, N. Carlini, and D. Wagner, "Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples," arXiv preprint arXiv:1802.00420, 2018.

[183] J. M. Zhang, M. Harman, L. Ma, and Y. Liu, "Machine learning testing: Survey, landscapes and horizons," IEEE Transactions on Software Engineering, 2020.

[184] S. Bhattacharyya, D. Cofer, D. Musliner, J. Mueller, and E. Engstrom, "Certification considerations for adaptive systems," in 2015 International Conference on Unmanned Aircraft Systems (ICUAS). IEEE, 2015, pp. 270–279.

[185] F. K. Došilović, M. Brčić, and N. Hlupić, "Explainable artificial intelligence: A survey," in 2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO). IEEE, 2018, pp. 0210–0215.

[186] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," Proceedings of the IEEE, vol. 103, no. 1, pp. 14–76, Jan. 2015.

JANI SUOMALAINEN received his M.Sc. (Tech.) degree in Information Technology from the Lappeenranta University of Technology, Finland in 2001 and Lic.Sc. (Tech.) degree on Telecommunications Software from the Aalto University, Finland in 2013. Since 2000 he has been with VTT Technical Research Centre of Finland in Espoo where he is a Senior Scientist. He is specialized on cybersecurity and has been involved in these topics in various international joint projects and customer projects. He has researched smart security applications, security interoperability, as well as developed ML-based threat detection and security situation awareness systems for software-defined mobile networks. He has also participated to the development of security architecture for 5G. Currently, he has been involved in both European and Finnish cooperation projects developing secure network slicing and ML-based security monitoring solutions for emerging cellular networks. His research interests include adaptive and learning security solutions for dynamic and heterogeneous network environments. He is a co-author of more than 30 scientific publications on network security.

ARTO JUHOLA received his M.Sc. (Electrical Engineering) degree from the Tampere University of Technology, Finland in 1987. He has a long background on the telecommunication industry and research. He has worked as a Development Engineer at Teleste Antenna Ltd 1984-89, and at Helsinki Telephone Association 1989-94. The work included transmission, fast data network, Internet and network management development. From 1995 he worked as a Senior Researcher at Helsinki Telephone Ltd.'s Research Center. His research focus was in Internet, middleware, active network, network management, and intelligent network research. In 2000 he worked as an Active Networking Group Leader at Elisa Communication. During 2001 - 2006 he was a Mobility Group Leader at VTT Technical Research Centre of Finland, and the project manager of the EU project CONTEXT, which researched active networks and context awareness. From 2006 onwards, he has been involved in various Internet security related projects, international/national joint, and industry contract, as a Senior Scientist at VTT. Recent involvement in 5G security includes the European cooperation projects SASER and SENDATE PLANETS where he participated in the application of ML to anomaly detection as a precursor to subsequent real-time risk estimation methods and ML assisted decision making for reactive security.

SHAHRIAR SHAHABUDDIN received his M.Sc. and Ph.D. degrees from the Centre for Wireless Communications, University of Oulu, Finland in 2012 and 2019, respectively, under the supervision of professor Markku Juntti. During Spring 2015, he worked at the Computer Systems Laboratory of Cornell University, NY, USA in Professor Christoph Studer's group. Shahriar received distinction in his M.Sc. degree and the best master's thesis award of the Department of Communications Engineering, University of Oulu in 2012. He received several scholarships and grants such as Nokia Foundation Scholarship, University of Oulu Scholarship Foundation Grant, Tauno Tönning Foundation Grant during his Ph.D. studies. Shahriar's research interest includes VLSI signal processing, MIMO detection and precoding, 5G and 6G security, and machine learning applications for wireless communications. Since 2017, Shahriar has been with Nokia, Finland as a SoC Specialist.

AARNE MÄMMELÄ received the degrees of M.Sc. (Tech.) and D.Sc. (Tech.) (both with honors) from the University of Oulu in 1983 and 1996, respectively. He was with the University of Oulu from 1982 to 1993. In 1993 he joined VTT Technical Research Centre of Finland in Oulu. Since 1996 he has been a Research Professor of digital signal processing in wireless communications. He has visited the University of Kaiserslautern in Germany in 1990-1991 and the University of Canterbury in New Zealand in 1996-1997. Since 2004 he has been also a Docent (equivalent to Adjunct Professor) at the University of Oulu. In 2014-2018 he was a Technical Editor of the IEEE Wireless Communications and in 2016-2018 a member of the Research Council of Natural Sciences and Engineering in the Academy of Finland. He has given lectures on research methodology at the University of Oulu for about 20 years, including the interdisciplinary holistic systems approach in addition to the conventional reductive or analytical approach. He has published tutorial and review papers on systems thinking in the IEEE Circuits and Systems Magazine and in the IEEE Access. His research interests are in intelligent adaptive, learning, and autonomous systems and resource efficiency in telecommunications.

IJAZ AHMAD is a Member of the IEEE. He received his MSc. (Tech.) degree in 2012 and D.Sc. (Tech.) degree from the University of Oulu in 2018, both in wireless communications. He is a research scientist at VTT Technical Research Centre of Finland. Ijaz is a co-author in over 35 publications including an edited book on 5G Security with Wiley Inc., and two IEEE Communication Surveys & Tutorials articles on SDN security and 5G security as the main author. Dr. Ijaz has received several awards including the Nokia Foundation, Tauno Tönning, and Jorma Ollila grant awards, and two IEEE best paper awards. He has been a visiting scientist at Aalto University, Finland in 2018. In 2019, he visited TU Vienna, Austria to work with Prof. Thilo Sauter as a visiting scientist. His research interests include the application of machine learning in wireless networks and 5G, 5G security, SDN, and security of machine learning techniques.