Security and Privacy in IoT Using Machine Learning and Blockchain: Threats and Countermeasures

NAZAR WAHEED and XIANGJIAN HE, University of Technology Sydney, Australia
MUHAMMAD IKRAM, Macquarie University, Australia
MUHAMMAD USMAN, University of South Wales, United Kingdom
SAAD SAJID HASHMI, Macquarie University, Australia
MUHAMMAD USMAN, Federation University, Australia

Security and privacy of users have become significant concerns due to the involvement of the Internet of Things (IoT) devices in numerous applications. Cyber threats are growing at an explosive pace making the existing security and privacy measures inadequate. Hence, everyone on the Internet is a product for hackers. Consequently, Machine Learning (ML) algorithms are used to produce accurate outputs from large complex databases, where the generated outputs can be used to predict and detect vulnerabilities in IoT-based systems. Furthermore, Blockchain (BC) techniques are becoming popular in modern IoT applications to solve security and privacy issues. Several studies have been conducted on either ML algorithms or BC techniques. However, these studies target either security or privacy issues using ML algorithms or BC techniques, thus posing a need for a combined survey on efforts made in recent years addressing both security and privacy issues using ML algorithms and BC techniques. In this article, we provide a summary of research efforts made in the past few years, from 2008 to 2019, addressing security and privacy issues using ML algorithms and BC techniques in the IoT domain. First, we discuss and categorize various security and privacy threats reported in the past 12 years in the IoT domain. We then classify the literature on security and privacy efforts based on ML algorithms and BC techniques in the IoT domain. Finally, we identify and illuminate several challenges and future research directions using ML algorithms and BC techniques to address security and privacy issues in the IoT domain.

CCS Concepts: • **Security and privacy** → *Security services*;

Additional Key Words and Phrases: Blockchain, cybersecurity, Internet of Things, machine learning

This research is supported by the Australian Government Research Training Program Scholarship.

Authors' addresses: N. Waheed and X. He (corresponding author), School of Electrical and Data Engineering, University of Technology Sydney, 2007, Sydney, New South Wales, Australia; emails: nazar.waheed@student.uts.edu.au, xiangjian.he@uts.edu.au; M. Ikram and S. S. Hashmi, Department of Computing, Macquarie University, 2109, North Ryde, New South Wales, Australia; emails: muhammad.ikram@mq.edu.au, saad.hashmi@hdr.mq.edu.au; M. Usman, School of Computing and Mathematics, University of South Wales, CF37 1DL, Pontypridd, Rhondda Cynon Taff, United Kingdom; email: muhammad.usman@southwales.ac.uk; M. Usman, School of Science, Engineering and IT, Federation University, 3350, Mt. Hellen, Victoria, Australia; email: muhammad.usmanskk@gmail.com.

0360-0300/2020/12-ART122 $15.00
https://doi.org/10.1145/3417987

ACM Reference format:
Nazar Waheed, Xiangjian He, Muhammad Ikram, Muhammad Usman, Saad Sajid Hashmi, and Muhammad Usman. 2020. Security and Privacy in IoT Using Machine Learning and Blockchain: Threats and Countermeasures. *ACM Comput. Surv.* 53, 6, Article 122 (December 2020), 37 pages.
https://doi.org/10.1145/3417987

1 INTRODUCTION

We have seen industries evolve from manufacturing just the *products* to building the *network of products* known as the *Internet of Things* (IoT), and eventually creating an intelligent *network of products* providing various, invaluable online service [27, 64]. As per Aksu et al. [7], two devices are connected to the Internet every 3 minutes. This connectivity and the exponential growth of IoT devices have resulted in an increased amount of network traffic. Due to this connectivity, challenges like security and privacy of user data and verification and authentication of devices, have arisen [118]. For example, hackers compromised one billion Yahoo accounts in 2013 [48]. In 2014, 145 million eBay users were under attack [99]. Following the increasing trend of attacks, in 2017, 143 million customers from Equifax had their personal information stolen [146]. Similarly, as reported in Ref. [75], a five billion dollar toy industry in 2017 had their 820,000 client accounts compromised. It also included over two million voice recordings, out of which a few were held for ransom. The recent cyber history is full of cybersecurity disasters, from massive data breaches to security flaws in billions of microchips and computer system lockdowns until a payment was made [45]. There are a plethora of security and privacy challenges for IoT devices, which are increasing every day. Hence, security and privacy in complex and resource-constrained IoT environments are big challenges and need to be tackled effectively.

The security challenges in IoT are increasing as the attacks are getting sophisticated day by day. Milosevic et al. [95] highlighted that powerful computing devices, e.g., desktop computers, might be able to detect malware using sophisticated resources. However, IoT devices have limited resources. Similarly, traditional cybersecurity systems and software are not efficient enough in detecting small attack variations or zero-day attacks [21], since both need to be updated regularly. Moreover, the updates are not available by the vendor in real time, making the network vulnerable. Machine Learning (ML) algorithms can be employed to improve IoT infrastructure (such as smart sensors and IoT gateways) [28], and also to improve the performance of cybersecurity systems [139]. Based on the existing knowledge of cyber-threats, these algorithms can analyze network traffic, update threat knowledge databases, and keep the underlying systems protected from new attacks [7, 137, 139]. Alongside using ML algorithms, the researchers have also started using revolutionary Blockchain (BC) technique to protect the underlying systems [37, 38, 77–79, 85, 105, 131, 145]. Although ML algorithms and BC techniques have been developed to deal with cyber threats in the IoT domain, combining these two is something new that needs to be explored.

Privacy goes hand-in-hand with security. Price et al. defined privacy as an application-dependent set of rules [104]. The authors elaborate that the rules on how the information can flow depend on the involved entities, processes, frequency, and motives to access data. There are many applications, such as wearable devices [7], Vehicular Area NETwork (VANET) [143], healthcare [147], and smart-home [31, 32, 120] that require providing security and protecting the privacy of personal information. For example, in a crowdsensing application like VANET, the network is dependent on the data collected from devices to make intelligent decisions on the latest traffic conditions. However, the users of devices might be hesitant to participate due to inadequate privacy-preserving mechanisms and related threats. Extensive research works based on ML algorithms and BC techniques [7, 37, 38, 77–79, 85, 105, 131, 137, 139, 145] have been conducted in the past few years to protect data on devices and preserve user's privacy.

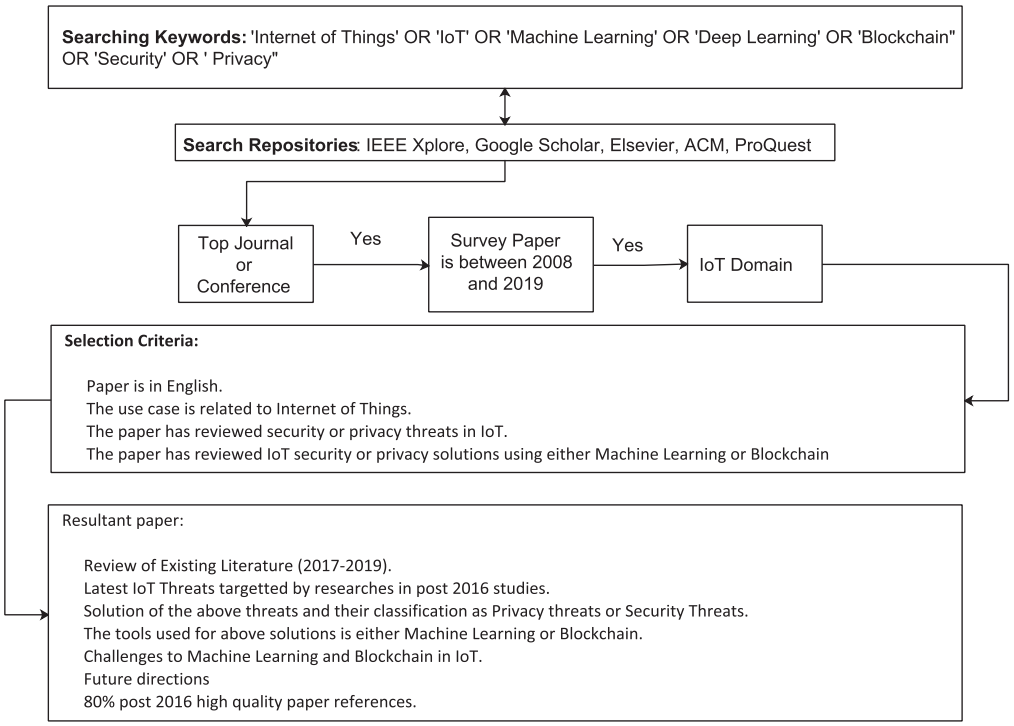


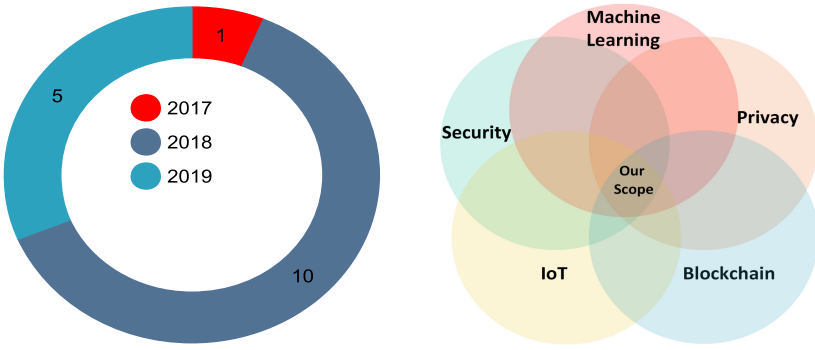
Fig. 1. Paper collection criteria.

Paper collection: Figure 1 depicts the strategy of selecting articles for this study. Initially, using the keywords and mentioned databases, the search was performed. The keywords such as IoT, Internet of Things, privacy, security, machine learning, and blockchain were utilized to download the latest articles from the top journals and conferences. In order to qualify for selection, an article must satisfy *all* of the following conditions: (i) published between 2008 and 2019 (inclusive); (ii) be a generic (not application specific) IoT survey paper; (iii) discussed security or privacy threats related to IoT; and (iv) covered ML and/or BC as a computing paradigm. The year-wise articles selection statistics are depicted in Figure 2(a).

Contributions of the article: This article provides a detailed review of ML algorithms and BC techniques employed to protect IoT applications from security and privacy attacks. Based on the review, we highlight that a combination of ML algorithms and BC techniques can offer more effective solutions to security and privacy challenges in the IoT environment. To the best of our knowledge, this is the first article that presents a review of security and privacy vulnerabilities in the IoT environment and their countermeasures based on ML algorithms and BC techniques. A road map of our article is depicted in Figure 3, while Figure 2(b) illustrates the scope of this survey article.

To cover the gaps in current literature (as summarized in Table 1), the major contributions of this article can be summarized as follows.

- We provide a generic classification of IoT threats reported in recent literature based upon security and privacy threats.
- We classify literature reviews on ML algorithms and BC techniques for IoT security and privacy, and highlight the research gaps in the existing literature reviews as in Tables 4–6.



(a) Year-wise distribution (b) Survey Scope

Fig. 2. (a) Year-wise statistics of the selected survey papers between 2008 and 2019 inclusive. It shows that most of the work only started recently. (b) Scope of our study highlighting the use of ML and BC techniques to address security and privacy issues in IoT domains.

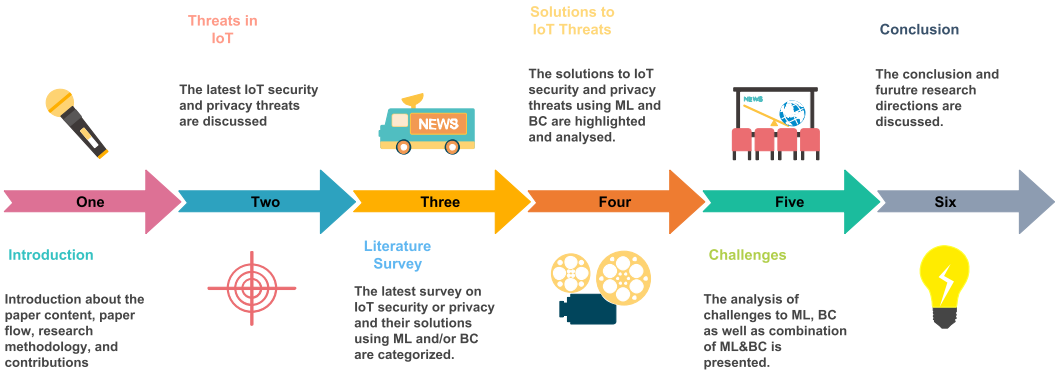


Fig. 3. Roadmap of our study.

—We provide a taxonomy of the latest security and privacy solutions in IoT using ML algorithms and BC techniques.
—We also identify and analyze the integration of ML algorithms with BC techniques to strengthen security and privacy in IoT.
—Finally, we highlight and discuss existing challenges to ML algorithms and BC techniques in IoT security and privacy with an attempt to suggest some future directions.

The rest of this article is organized as follows. In Section 2, the classification of well-known IoT threats is presented. In Section 3, we categorize literature reviews on IoT security and privacy using ML algorithms and BC techniques. Section 4 presents the latest solutions to IoT security and privacy threats, whereas research challenges for techniques based on ML and BC to solve security and privacy issues are presented in Section 5. Finally, in Section 6, we conclude by presenting the gaps with some future directions.

2 THREATS IN IOT

IoT refers to a large number of heterogeneous sensing devices communicating with each other, either in a LAN or over the Internet [58]. IoT threats are different from conventional networks, significantly due to the available resources of end devices [67]. IoT devices have limited

Table 1. Contributions and Gaps of All Published Survey Papers from 2017 to 2019

Authors [Ref.]	Year	IoT Security	IoT Privacy	Machine Learning	Blockchain
Kshetri et al. [75]	2017	X	✓	X	✓
Banerjee et al. [14]	2018	✓	X	X	✓
Restuccia et al. [106]	2018	✓	X	✓	X
Sharmeen et al. [115]	2018	✓	X	✓	X
Xiao et al. [138]	2018	X	✓	✓	X
Khan et al. [73]	2018	✓	X	X	✓
Reyna et al. [107]	2018	✓	X	X	✓
Panarello et al. [60]	2018	✓	X	X	✓
Kumar et al. [76]	2018	✓	✓	X	✓
Kouicem et al. [74]	2018	✓	✓	X	✓
Zhu et al. [148]	2018	X	✓	X	✓
Chaabouni et al. [21]	2019	✓	X	✓	X
Hassija et al. [57]	2019	✓	X	✓	✓
Costa et al. [26]	2019	✓	X	✓	X
Wang et al. [134]	2019	✓	X	X	✓
Ali et al. [9]	2019	✓	✓	X	✓
This Survey	2020	✓	✓	✓	✓

memory and computational power, whereas the conventional Internet comprises powerful servers and computers with plentiful resources. Due to this, a traditional network can be secured by multi-factor security layers and complex protocols, which is what a real-time IoT system cannot afford. In contrast to traditional networks, IoT devices use less secure wireless communication media such as LoRa, ZigBee, 802.15.4, and 802.11a/b/n/g/p. Lastly, due to application-specific functionality and lack of common OS, IoT devices have different data contents and formats, making it challenging to develop a standard security protocol [89]. All these limitations make IoT prone to multiple security and privacy threats, thus opening venues for various types of attacks.

The probability of an attack in a network increases with the network size. Therefore, the IoT network has more vulnerabilities than a traditional network, for example, a company office. Additionally, IoT devices communicating with each other are usually multi-vendor devices with different standards and protocols. The communication between such devices is a challenge, which requires a trusted third party to act as a bridge [17]. Moreover, several studies have raised the concern of regular software updates to billions of smart devices [42, 77].

The computational resources of an IoT device are limited, so the capabilities of dealing with advanced threats are degraded. To summarize, IoT vulnerabilities can be categorized as *specific* and *common*. For example, vulnerabilities like *battery-drainage attack*, *standardization*, and *lack of trust* are *specific* to IoT devices, and Internet-inherited vulnerabilities can be regarded as *common vulnerabilities*. Several IoT threats and their categorization have been introduced in the past [19, 96, 106, 138, 139]. We discuss the most common threats in IoT reported in the past decade and attempt to classify them into security and privacy categories.

2.1 Security Threats

The fundamental concepts of security and privacy revolve around the triad of **C**onfidentiality of the data, **I**ntegrity of data, and **A**vailability of the network (CIA) [18, 102, 142]. In IoT, data can be anything, for example, a user's identity information, packets sent from a surveillance camera to a destination server, a command given by a user to its car using a key-fob, or a multimedia

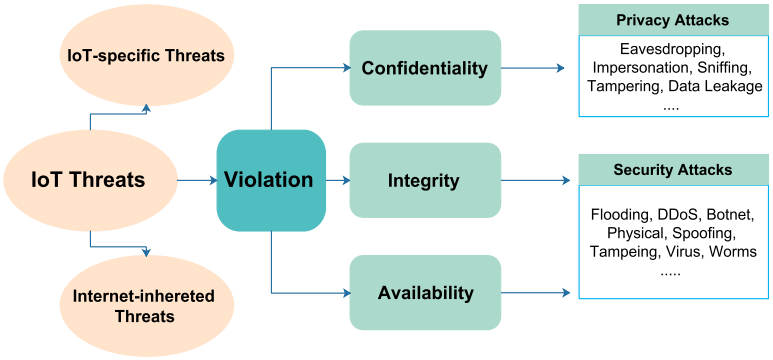


Fig. 4. Types of IoT threats may violate any of the CIA triad; *integrity* & *availability* are associated with security attacks, while *confidentiality* compromise is known as a privacy attack.

conversation between two people. Any unauthorized disclosure of data may result in a violation of either confidentiality, integrity, or availability. If a threat is impacting confidentiality, it is a privacy threat. The security threats affect both data integrity and network availability. Figure 4 depicts different classes of security and privacy threats in IoT domains.

2.1.1 Denial of Service. Denial of Service (DoS) has the most straightforward implementation among all the security attacks comparatively. Furthermore, the ever-growing number of IoT devices with weak security features has made DoS a favorite tool for attackers. The core objective of a DoS attack is to ingest the network with invalid requests, resulting in exhausting network resources, such as bandwidth consumption. As a result, the services are unavailable to genuine users. Distributed DoS (DDoS) is an advanced version of the DoS attack, where multiple sources attack a single target making it more difficult to trace and avoid the attack [1, 30, 113, 123, 124, 128]. Although there are different types of DDoS attacks, they all have the same objective. Few variants of DDoS attacks are SYN flooding [68] (in which an attacker sends successions of SYN requests to a target), Internet Control Message Protocol (ICMP) attacks [34] (in which a large number of ICMP packets are broadcasted using the victim's spoofed IP), crossfire attacks [108] (using a complex and massively large-scale botnet for attack execution), and User Datagram Protocol (UDP) flooding attacks [72] (sending a large number of UDP packets to random ports on a remote victim). Botnet attack [97] is a type of DDoS attack in an IoT network. The botnet is a network of IoT nodes (devices) that are compromised to launch an attack on a specific target, for example, a bank server. Botnet attack can be executed on different protocols, particularly Message Queuing Telemetry Transport (MQTT), Domain Name Server (DNS), and Hypertext Transfer Protocol (HTTP), as briefed in Ref. [97]. Several techniques to prevent DoS in the IoT environment are suggested. Diro et al. [30] have utilized the self-learning characteristics of Deep Learning (DL) methods to detect an attack in the fog-to-things environment. In another study, Abeshu et al. [1] suggested controlling the DDoS attack by employing distributed DL on fog computing. Intrusion Detection System (IDS) by Tan et al. in Refs [123] and [124] is a series of research efforts to mitigate DDoS attacks using modern ML and DL algorithms. Sharma et al. [113] and Tselois et al. [128] pointed out the issues of flooding in Software Defined Networks (SDN). The study highlighted that the SDN's top layer was prone to brute force attacks due to the lack of authentication in the plain-text Transmission Control Protocol (TCP) channel.

2.1.2 Man-in-the-middle. Man-in-the-middle (MiTM) attacks are one of the oldest attacks in the cyber world [122]. Spoofing and impersonation can be categorized as MiTM attacks. For example, a node X intending to communicate with destination B might be communicating with the MiTM

attacker, who is impersonating to be destination B. Similarly, in Secure Sockets Layer (SSL) striping, an attacker can capitalize on such attacks to connect themselves with the server using an HTTPS connection, but with the target on an unsecured HTTP connection. Recently, many studies have focused on improving the security against MiTM attacks [4, 11, 24, 125, 132]. Ahmad et al. [4] discussed a healthcare scenario where a patient gets an insulin dosage automatically. Such an application is vulnerable to MiTM attack, which can prove fatal. For instance, Tang et al. [125] identified vulnerabilities in mobile apps' network service libraries, which can potentially expose apps' traffic to MiTM attacks. Similarly, in line with the impersonation attacks, Chatterjee et al. [24] highlighted existing methods of authentication in wireless mobile devices that used a secret key. This key was saved in non-volatile memory and used for digital signatures or hash-based encryption. Apart from being vulnerable, this technique was power inefficient. Similarly, the most recent and widely used IoT protocol, called OAuth 2.0, suffers from cross-site-recovery-forgery (CSRF) attacks. The OAuth protocol manually authenticates devices, which is a time-consuming process. Another study by Wang et al. [132] mentioned physical-layer security vulnerability in wireless authentication. They argued that the existing hypothesis test to compare radio channel information with channel record of *Alice* to detect a spoofer *Eve* in wireless networks is often unavailable, particularly in dynamic networks.

2.1.3 Malware. *Malware* is an abbreviation of **mal**icious soft**ware**. Over the last couple of years, the number of IoT devices is growing along with frequent IoT software patches, which may be leveraged by an attacker to install malware on a device and perform malicious activities. Malware is generally understood to exist as a virus, spyware, worm, trojan horse, rootkit, or malvertising [89, 144]. Smart home products, healthcare devices, and vehicular sensors are a few examples that can be compromised. Azmoodeh et al. [13] studied malware on the Internet of Battlefield Things (IoBT). Such attackers are usually state-sponsored, better-resourced, and professionally-trained. Aonzo et al. [12], Feng et al. [40], and Wei et al. [135] attempted to defend resource-constrained android devices from malware attacks by using different supervised ML algorithms. Studies in Refs [52], [77], and [115] provided a detailed analysis of malware detection and highlighted several security loopholes in the Android platform, especially on the application layer, which has applications with several types of components.

2.2 Privacy Threats

In addition to security threats, IoT users and their data are prone to privacy attacks, such as sniffing, de-anonymization, and inference attacks. In any case, the impact is on the confidentiality of data, where data can be at rest or in motion. In this section, we discuss various privacy attacks.

2.2.1 MiTM. We believe that MiTM attacks can be classified into Active MiTM Attacks (AMA) and Passive MiTM Attacks (PMA). The PMA passively listens to data transfer between two devices. Although the PMA violate privacy, they do not alter the data. An attacker with access to a device can silently observe for months before attempting the attack. With the growing number of cameras in IoT devices like toys, smartphones, and wristwatches, the impact of PMA, for example, eavesdropping and sniffing, is immense. On the other hand, the AMA are actively involved in abusing the data acquired by either interacting with a user pretending to be someone else, for example, impersonation, or accessing a profile without consent, for example, authorization attack.

2.2.2 Data Privacy. Similar to MiTM attacks, the data privacy attacks can be classified into *Active Data Privacy Attacks* (ADPA) and *Passive Data Privacy Attacks* (PDPA). Data privacy is related to data leakage [130], data tampering, identity theft, and re-identification [8]. The re-identification attacks are also known as inference attacks and are based on de-anonymization attacks, location

Table 2. Security Threats in IoT

Threat	Impact	Attack	Type	Layer of Impact	Solution
Security	Availability	DoS	Flooding	Physical, MAC	Multiple
			DDoS	Physical, MAC	Multiple
			Botnet	Physical, MAC	Multiple
		Physical	Damage	Physical	Physical Security
			Environmental	Physical	Shielding
			Power Loss	Physical	Uninterrupted power
			Hardware Failure	Physical	Backup
			Tampering	Physical	Physical Security
	Integrity	MiTM	Sybil Attack	Physical, MAC, Network	Code attestation, radio resources testing, key pool
			Spoofing message tamper	Network	anti-spoofing software
		Malware	Injection	Application	
			Virus	Application	
			Worms	Application	

Table 3. Privacy Threats in IoT

Threat	Impact	Attack	Type	Layer of Impact	Solution
Privacy	Confidentiality	MiTM	Eavesdropping	Network	Encryption
			Impersonation	Network	Encryption
			Sniffing	Network	Encryption
			Authroization	Application	Access Control
		Data Privacy	Data Leakage	Multiple	
			Re-identification	Multiple	Data suppresion, generalization, noise addition
			Data tampering	Multiple	Anonymization
			Identity Theft	Multiple	Anonymization
		Others	Poodle	Transport	Use TLSv1.2
			Heartbleed	Transport	
			Freak	Transport	Turnoff export ciphersuit options in browser

detection, and aggregation of information [8]. In these attacks, hackers' main goal is to gather data from multiple sources and reveal the targets' identities. Some attackers may use the collected data to impersonate an individual target [138]. Any attack that alters data, such as data tampering, can be classified as ADPA, while the re-identification and data leakage are examples of PDPA.

A comparison between various security and privacy threats, their types, their impacts, and possible solutions are summarized in Tables 2 and 3.

3 LITERATURE SURVEY

This section provides an existing literature review and categorizes the efforts done based on ML algorithms and BC techniques to address IoT security and privacy issues. This section is divided into two subsections, i.e., ML algorithms and BC techniques.

3.1 Existing Review Papers Using Machine Learning Algorithms as a Solution

Hackers are getting sophisticated with the evolving technology, making traditional methods of attack-prevention cumbersome. The defense becomes more challenging for a resource-constraint IoT device. To help in detecting these attacks, one of the widely used tools is ML algorithms. ML can be defined as the ability to deduce knowledge from data and adjust the output of an ML model based on that acquired knowledge [58]. ML makes machines smart enough by learning from their past results and refining them to achieve improved results [129]. Several ML algorithms have proven extremely helpful in mitigating security as well as privacy attacks. In the following sections, we discuss these approaches in detail.

3.1.1 Security Efforts. The technology has improved data communication and networking techniques over the Internet. We now have state-of-the-art software-based configurable devices, SDN, that can be customized to meet a customer's needs. In this scenario, Restuccia et al. [106] attempted to present the taxonomy of existing IoT security threats and their solutions in SDN using the ML algorithms. They also suggested that since the main task of an IoT system is to collect data from IoT devices, it is feasible to divide the data collection process into three steps, namely IoT authentication, IoT wireless networking, and IoT data aggregation and validation. The study gave a brief review of ML algorithms used to mitigate the security attacks, e.g., to detect cross-layer malicious attacks, Bayesian learning is used, and to assess the validity of data, neural networks are used. However, the study lacks an in-depth analysis of the rest of the ML algorithms.

Sharmeen et al. [115] aimed to assist application developers in using Application Program Interfaces (APIs) safely during the development of applications for Industrial IoT networks. To detect malware, the authors suggested that the ML model could be trained by using three types of features including static, dynamic, and hybrid. A detailed analysis of each feature type is done using performance metrics of a dataset, features extraction technique, features selection criteria, accuracy, and detection method. Several detection methods for each feature set were analyzed, but those commonly used were Random Forest (RF), Support Vector Machine (SVM), k-Nearest Neighbors (kNN), J48, and Naive Bayes (NB). Sharmeen et al. [115] concluded that hybrid analysis offered flexibility in choosing both the static and dynamic features to improve accuracy in the detection process. However, this article is limited to one application (android device) and one security threat (malware).

Costa et al. [26] selected papers between 2015 to 2018 and claimed that no work has presented an in-depth view of the application of ML in the context of IoT intrusion detection. The study reviewed the latest as well as traditional ML-based algorithms to improve IoT security. They also presented the most commonly used datasets and methodologies employed in the paper related to IoT security. The paper, however, has not reviewed the latest IoT security or privacy threats.

Similarly, Chaabouni et al. [21] also focused on the IoT-based network intrusion detection systems. The authors presented IoT architecture and layer-wise attacks, and classified them by layers (*perception layer*, *network layer*, and *application layer*) as well as design challenges (such as *heterogeneity, mobility, trust and privacy, resource constraints, connectivity, and data interchange*). The traditional mechanisms to protect IoT were described, and the study focused on Anomaly and Hybrid Network IDS (ANIDS) for IoT systems. A detailed comparison of traditional NIDS for IoT systems architecture, detection methodologies, and experimental results was provided. The study further presented how the learning-based NIDS for IoT could overcome the challenges faced by their equivalent traditional IoT systems. Finally, top IoT NIDS proposals were compared with a focus on ML algorithms.

All of the above papers, as depicted in Table 4, are limited to security threats with a focus on ML as a tool in solving the security issues. Our article, as depicted in Figure 2(a), covers a broader scope—addressing security and privacy issues in IoT domains using ML and BC.

Table 4. List of Survey Papers on IoT Security Leveraging Machine Learning Algorithms

Ref.	Security Threats	Proposed solution(s)
Restuccia et al. [106]	DoS, MiTM	A taxonomy and survey of IoT security research and their ML-based solutions
Sharmeen et al. [115]	Malware	Analysis of malware detection for Android mobile
Chaabouni et al. [21]	Multiple	A detailed analysis of traditional and ML-based NIDS for IoT
Costa et al. [26]	Multiple	In-depth review of ML applications in the context of IoT intrusion detection

3.1.2 Privacy Efforts. Machine learning extracts useful information from the raw data, while privacy is preserved by concealing the information [65]. According to Al-Rubaie et al. [8], the ML system has three modules: (i) input, (ii) computation, and (iii) output. The study further claimed that privacy could only be preserved if all three modules were under the ownership of a single entity. Nowadays, the data is collected worldwide by billions of IoT devices such as smart-phones, health monitoring sensors, speed cameras, and temperature sensors; hence, a *single-ownership* condition cannot be maintained. This issue spurred interest in researchers to work toward proposing newer and improved privacy-preserving ML algorithms. For instance, the lack of privacy protection mechanisms in a VANET environment was raised by Zhang et al. [143]. In VANET, vehicle nodes tend to learn collaboratively, raising privacy concerns, where a malicious node can obtain sensitive data by inferring from the observed data. A single node has limited computational and memory resources. The solution was presented by using collaborative IDS with distributed ML algorithms and resolving the privacy issues by proposing the concepts of dynamic differential privacy to protect the privacy of a training dataset.

People traffic monitoring systems and healthcare services are two of the most common IoT sensing technologies, which need continuous improvements. The most effective and useful data for such applications is directly collected from the users through Mobile CrowdSensing (MCS). Xiao et al. [138] reviewed the privacy threats involved in MCS, where the information of interest is extracted, and the participants upload sensing reports of their surroundings to the MCS server. This information-sharing poses significant privacy threats to the participants and the MCS server. The system is prone to privacy leakage (which is related to user's personal information), faked sensing attacks (sending fake reports to the server to reduce the sensing efforts), and advanced persistent threats (causing privacy leakage over an extended period). The survey suggested Deep Neural Network (DNN) and Convolutional Neural Network (CNN) for privacy protection, and Deep Belief Network (DBN) and Deep Q-Network (DQN) for counter-measuring faked sensing. However, the review was limited to only one application (MCS).

3.2 Existing Review Papers Using Blockchain as a Solution

Blockchain, often confused by some as a synonym to Bitcoin, is the technology behind this infamous crypto-currency. It is a distributed ledger that stores the data in blocks. These blocks are in order and linked with each other cryptographically forming a chain in a way that makes it computationally infeasible to alter the data in a particular block [23]. This mechanism ensures immutation, decentralization, fault-tolerance, transparency, verifiability, audit-ability, and trust [25, 60]. There is no single consensus on the types of BC, but, most commonly, they are public, private, and consortium. Public or permission-less BC is open to everyone, so anyone can access them [42]. On the other hand, private/permissioned blockchains are controlled by one or few; hence, not everyone can access them. The transactions here are faster and only the selected few are

Table 5. List of Survey Papers on IoT Security Leveraging Blockchain Techniques

Ref.	Year	Security Threats	Comments
Banerjee et al. [14]	2018	Several	Classified post-2016 literature & discusses BC-based solutions
Khan et al. [73]	2018	Key management	Categorization of threats & their BC-based solutions access control were presented
Reyna et al. [107]	2018	DoS	Challenges & Analysis of BC in IoT devices were mentioned
Panarello et al. [60]	2018	Multiple	Comprehensive BC-IoT integrated security challenges and emerging solutions were discussed
Kumar et al. [76]	2018	MiTM	How BC can be a solution for IoT security issues is discussed
Kouicem et al. [74]	2018	Multiple	Provided BC-based solution to attain a "trio" of *anonymity*, *unlinkability*, and *intractability*
Wang et al. [134]	2019	Multiple	IoT layer-wise attacks discussed; BC-based security solutions for IoT applications were discussed
Ali et al. [9]	2019	DoS	Reviewed latest proposed BC-based IoT security solutions
Hassija et al. [57]	2019	Multiple	A detailed survey of existing IoT security solutions is presented

Table 6. List of Survey Papers on IoT Privacy Leveraging Blockchain Techniques

Ref.	Year	Privacy Threats	Comments
Kshetri et al. [75]	2017	Identity management	Highlighted how BC is superior to the current IoT ecosystem
Kumar et al. [76]	2018	Spoofing, authentication	Presented IoT security and privacy issues and how BC can be a solution
Kouicem et al. [74]	2018	Data Privacy	Provided BC-based solution to attain a "trio" of *anonymity*, *unlinkability*, and *intractability*
Zhu et al. [148]	2019	Data Privacy	Highlighted challenges in traditional IdM systems and reviewed their BC-based solutions
Hassan et al. [56]	2019	Multiple	Comprehensively surveyed privacy preservation techniques of BC-based IoT systems from application and implementation
Ali et al. [9]	2019	Data Privacy, MiTM	Reviewed latest proposed BC-based IoT privacy solutions

authorized to approve a transaction, hence, reaching a consensus. Several reviews and survey papers [14, 25, 33, 42, 60, 73, 75, 80, 89, 94, 107, 127, 141] are published to highlight the importance of the BC techniques and could be a good source for those who are interested to read more about BC in detail. A detailed comparison of current work is shown in Tables 5 and 6. Most of these works discussed either security or privacy issues. In this section, we present the current literature reviews on achieving security and privacy in IoT using BC techniques as a tool.

3.2.1 Security Efforts. Security has been the prime focus of attention for any IoT use cases. Lots of work based on BC techniques have emerged to solve security issues in the IoT domain. A study on IoT security was presented by Banerjee et al. [14], which is classified into security techniques such as intrusion detection and prevention system (IDPS), collaborative security, and predictive security. Furthermore, IDPS are classified by approaches, network structure, and applications. After that, collaborative security and predictive security are discussed in detail. In the same study, collaborative security techniques are classified by network structures and applications. Sequel to this study, the integrity of existing IoT datasets is highlighted, and the authors suggested that a BC-based standard should be developed to ensure integrity in the shared datasets.

In another study by Khan et al. [73], security issues related to key management, access control, and trust management in IoT are discussed. Khan et al. [73] categorized the security threats into IoT

layers and presented their BC-based solutions. The IoT security issues were classified as *low-level*, *intermediate-level*, and *high-level* security issues. Khan et al. [73] believes that jamming adversaries, insecure initialization, spoofing, vulnerable physical interface, and sleep deprivation attacks are the low-level security issues. Whereas, replay, RPL routing attacks, sinkhole, Sybil attack on intermediate layers, transport-level end-to-end security, session establishment, and authentication are intermediate-level security issues. The high-level security issues are insecure interfaces, CoAP security with Internet, vulnerable software, and middleware security. The study then provided a comprehensive mapping of all the above problems with the affected layers of IoT architecture and proposed solutions for each one of them. In the end, the authors discussed how BC techniques could be used to address and solve some of the most pertaining IoT security problems. This survey highlighted the security risks involved in each IoT layer but lacked the discussion of providing BC-based solutions for these security threats.

Similarly, Reyna et al. in Ref. [107] analyzed how BC techniques could potentially improve the security (data reliability) in the IoT. The study mentioned security threats as one of the challenges for BC techniques. The security threats mentioned in the study were majority attacks, double-spend attacks, and DoS attacks. The study also provided highlights about the integration of IoT with BC techniques, BC applications, and BC platforms. However, the study did not cover several other security attacks related to IoT, which was a limitation of this survey.

On the other hand Panarello et al. [60] comprehensively reviewed BC consensus protocols in addition to security challenges and recent developments in IoT and BC integration. The past literature was categorized based on application areas, which were supported by an extensive survey of the latest BC-based solutions.

Kumar et al. [76] presented a brief overview of issues and challenges in IoT security, such as spoofing and false authentication. Some of the advantages of BC for large scale IoT systems are tamper-proof data, trusted and reliable communication, robustness, and distributed and delegated data sharing. Sequel to that, the study has discussed the application-wise BC-based IoT challenges.

The authors of Ref. [74] highlighted security issues in IoT and provided their BC-based solutions. The study first highlighted the IoT security requirements and its challenges in six different application domains, like smart cities, healthcare, smart grids, transport, smart homes, and manufacturing. The authors comprehensively discussed the taxonomy of IoT security solutions such as confidentiality and availability. They also investigated the analysis of techniques that were suitable for each IoT application.

Wang et al. [134] highlighted the limitation of IoT security and provided comprehensive security analysis on end devices, communication channels, network protocols, sensory data, DoS attack, and software attacks. After presenting the existing BC technologies, the application of BC for IoT and their challenges were discussed. The study also briefly discussed the security of IoT applications using BC.

The potential benefits and motivations for developing a BC-based IoT framework are *resilience, adaptability, fault tolerance, security and privacy, trust*, and *reduced maintenance cost*. [9] The study mentioned that the centralized IoT model is prone to DDoS attacks. Moreover, due to its architecture, it has a single point of failure, which is a threat to the availability of IoT services. Current IoT security solutions are centralized because they involve trusting in third party security services, which bring in data integrity issues. Ali et al. highlighted how all of these issues can be solved by using BC-based IoT security solutions.

A comparison of IT and IoT security, followed by a comprehensive classification of IoT applications and their security and privacy issues were discussed by Hassija et al. [57]. The study even went ahead and discussed various possible security threats in IoT applications for four layers, i.e., (i) sensing, (ii) network, (iii) middle-ware, and (iv) application. In the recommendations to improve

the IoT security, BC was also mentioned as one of the solutions. This article is probably the closest to our work; however, it is security-biased and does not focus on the IoT privacy issues in detail.

A table of the existing surveys focused on security using BC techniques for IoT applications is compiled in Table 5.

3.2.2 Privacy Efforts. Previous studies, such as Ref. [43], used strong cryptographic measures to protect against malicious third parties and provided accountable access to IoT. However, they did not use either ML or BC as one of their tools. Kshetri et al. [75] highlighted how BC techniques can offer better privacy-preserving solutions as compared to a traditional network for cloud-based services. It also highlighted the superiority of BC in identity management and the provision of access control. The study demonstrated how an attack on the IoT network could be contained using BC techniques. However, a comprehensive privacy-preserving IoT threat model using BC techniques was missing in this literature.

Kumar et al. [76] presented a brief overview of issues and challenges in IoT privacy, such as data sharing. Sequel to that, the work suggested the BC-based solutions to these challenges and discussed several application areas for BC implementation. Although the work discussed challenges to BC in the IoT application, it lacks a comprehensive discussion on the latest IoT security and privacy threats.

The authors of Ref. [74] highlighted privacy issues in IoT and provided their BC-based solutions. The main goal of privacy-preserving techniques was to attain a "trio" of *anonymity*, *unlinkability*, and *intractability*. The main security services, for example, *confidentiality*, *privacy*, and *availability*, were addressed based on traditional cryptographic approaches. The study addressed the issues of data-sharing, data privacy, and user's behavior in IoT, and discussed their solutions, for example, data tagging, zero-knowledge proof, pseudonyms, and k-anonymity model.

Zhu et al. [148] highlighted privacy vulnerabilities in a traditional Identity Management (IdM) system, especially due to their centralized architecture, and reliability on the so-called trusted third parties. These vulnerabilities may result in several privacy attacks such as phishing and data leakage. The authors argued that traditional IdM systems can not be directly transplanted to IoT environments due to some native IoT characteristics such as scalability, mobility, and compatibility. Sequel to that, the study highlighted the privacy challenges in traditional IdM systems and reviewed their BC-based solutions.

Hassan et al. [56] provided a detailed overview of privacy issues in BC-based IoT systems. The privacy attacks related to BC-based IoT networks such as *Address reuse, Deanonymization, Sybil attack, Message Spoofing*, and *Linking attacks* were highlighted. The work also discussed the implementation of the five most popular privacy preservation strategies (*Encryption, Smart Contract, Anonymization, Mixing,* and *Differential Privacy*) within BC-based applications.

Ali et al. [9] reviewed the IoT privacy issues and their latest BC-based solutions. They raised the privacy concerns in a centralized IoT model such as data privacy and data confidentiality. The existing centralized privacy solutions such as *using a privacy broker*, *using group signatures*, applying *k-anonymity*, and pseudonyms were all heavily dependent on third parties for their services. To counter these issues, the study offered a comprehensive review of the BC-based IoT privacy solutions.

4 SOLUTIONS TO IOT THREATS

Since the inception of the first virus (Creeper) in 1970 until the hack of Whatsapp on May 15, 2019 [100, 117] and later, security specialists have mitigated zero-day security or privacy threats [62, 123, 144]. Regarding this, several solutions have been proposed to mitigate security and privacy issues. However, in this section, we focus on the recent literature proposing secure and privacy-preserving

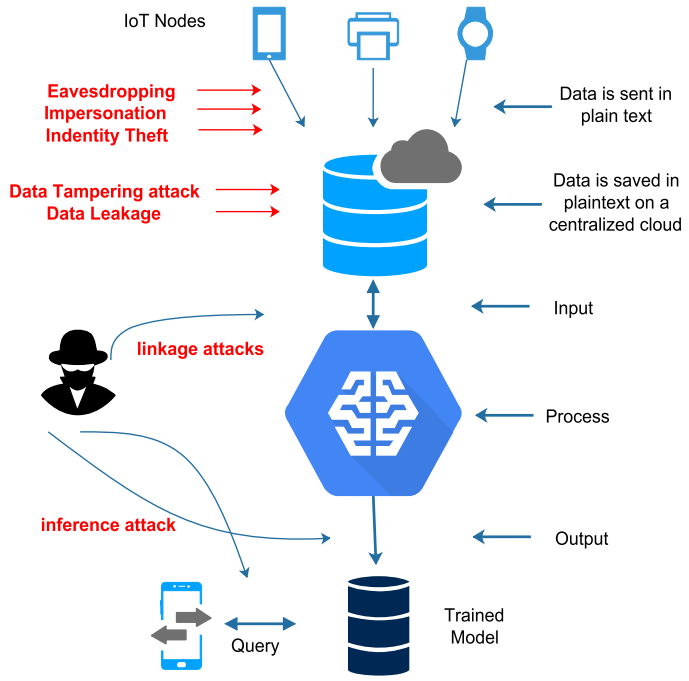


Fig. 5. An illustration of ML threat model for IoT: An ML model is prone to several attacks at either (i) *input*, (ii) *process*, or (iii) *output* stages.

techniques for the IoT domain. We discuss the solutions offered by first using ML algorithms as a tool, *then* by utilizing BC techniques, and *finally* by the fusion of both.

4.1 Existing Solutions Using Machine Learning Algorithms

ML is used as a data processing pipeline in any framework. For example, data traffic entering a network can be analyzed by an ML model to make an informed decision. The main components of the ML threat model for IoT are shown in Figure 5. Additionally, the figure gives an overview of target points, such as input and output, for an attacker. The input data from source to IoT nodes, and IoT nodes to ML model can experience exploratory or poisoning attacks. At the output, integrity and inversion attacks are possible [84]. Therefore, for a whole system to be completely immune to attacks, it must be secured as well as privacy-preserved.

4.1.1 Security Efforts. Several security solutions have been proposed using ML algorithms as a tool, as shown in Table 7. To deal with the flooding attacks, Diro et al. [30] argued that fog-computing reduced the risk of eavesdropping and MiTM attacks by restricting the communication to the proximity of IoT devices. Capitalizing on this idea, they used the Long Short Term Memory (LSTM) algorithm in their model as it can remember the older data. For binary classification, they compared their results with LR using ISCX2012 dataset, which had 440,991 normal traffic instances and 71,617 DoS attack instances. The DL model LSTM took considerably more time to train than LR, but its accuracy was 9% better. The second dataset used was Aegean Wi-Fi Intrusion Dataset (AWID) from Ref. [20], and consists of normal traffic instances (1,633,190 training and 530,785 tests), injection attack instances (65,379 training and 16,682 tests), flooding attack instances (94,848 training and 8,097 testings), and impersonation attack instances (48,522 training and 20,079 testings). After comparing LSTM against softmax for multi-class classification, the resultant accuracy obtained was 14% improved.

Table 7. Existing IoT Security Solutions Using Machine Learning Algorithms * Dataset notation in 'Dataset' column. I: ISCX2012, A: AWID, N: NSL-KDD, K: KDDCUP99, U: UNSW-NB15, NB: NIMS botnet, P: Private, AWI: Aegan WiFi Intrusion, Ab: AbdroZoo, D: Drebin, C: CTU-13, Ky: Kyoto 2006+

Ref.	Threat	Type of Threat	IoT Use case	Algo used	Feature Extraction	Feature Selection	Dataset	Accuracy
Diro et al. [30]	DoS	Flooding	Fog	LSTM	-	-	I, A	I (99.91), A (98.22)
Abeshu et al. [1]	DoS	Flooding	Fog	Softmax	SAE	-	N	99.2
Tan et al. [123]	DoS	Flooding	NIDS	TAB	MCA	Norm.	K	normalized 99.95
Tan et al. [124]	DoS	Flooding	CV	EMD	MCA	PCA	K, I	K (99.95), I(90.12)
Moustafa et al. [97]	Botnet	Flooding	IoT	Adaboost	CC	-	U, NB	U(99.54)
Ahmad et al. [4]	MiTM	Impersonation	Healthcare	LSTM RNN	NG	-	P	-
Aminanto et al. [11]	MiTM	Impersonation	WiFi	ANN	D-FES	-	AWI	99.92
Chatterjee et al. [24]	MiTM	Impersonation	RF Comm	ANN	-	-	P	99.9
Azmoodeh et al. [13]	Malware	Code Ijnection	IoBT	DCN	OpCodes	IG	P	98.37
Aonzo et al. [12]	Malware	Malware	Android	-	Static Analysis Technique	Manual	P	98.9
Wei et al. [135]	Malware	Malware	Android	NB, C4.5, kNN	Dynamic Analysis technique	NA	P	-
Feng et al. [40]	Malware	Malware	Android	ensemble + LR	Manual	Chi-Square	Ab, D	98.18
Wang et al. [133]	Malware	Malware	Android	ensemble	String + structural	ensemble	Multi-sources	98.4
Maimo et al. [88]	Anomaly	Anamoly	5G	LSTM	Weighted Loss	ASD (DBN+SAE)	C	-
Niyaz et al. [63]	Anomaly	Anamoly	NIDS	Softmax	SAE using Backpropogation	-	Ky	2- 88.39, 5- 79.10
Ambusaidi et al. [10]	Anomaly	Anomaly	NIDS	LSSVM	MMIFS	FMIS	K, N, Ky	K 99.95,I 90.12
Zhou et al. [146]	Dataset	Multiple	IoT	DFEL	-	-	N, U	>98.5
Prabavathy et al. [101]	Dataset	Multiple	Fog	OS-ELM	-	-	N	97.36

In a similar study, Abeshu and Chilamkurti highlighted that the resource constraints of an IoT device made it a potential threat to DoS attacks [1]. Classic ML algorithms are less accurate and less scalable for cyber-attack detection in a massively distributed network such as IoT. Such a massive amount of data produced by billions of IoT devices enables the DL models to learn better than the shallow algorithms. The authors of Ref. [1] argued that most of the employed DL architectures had used pre-training for feature extraction, which could detect anomalies and thus reduced the workload of a network administrator. However, their work was focused on distributed DL through parameters and model exchange for the applications of fog computing. Fog computing reduced the load of computing power and storage space from the IoT devices. It is, therefore, the ideal spot where an intrusion can be detected. The existing Stochastic Gradient Descent (SGD) for fog-to-things computing needs parallel computing. Thus, the centralized SGD will choke due to the massive amount of data in IoT. Therefore, the study proposed a distributed DL-driven IDS using NSL-KDD dataset, where the stacked auto-encoder (SAE) was used for feature extraction, and soft-max regression (SMR) was used for the classification. Their study proved that the SAE as a DL worked better than traditional shallow models in terms of accuracy (99.27%), FAR and DR. Both Diro et al. [30] and Abeshu et al. [1] proved that the DL algorithms performed better than shallow ML models.

As a first attempt to DoS detection, Tan et al. [123] used triangle-area-based technique to speed up the feature extraction in Multivariate Correlation Analysis (MCA). Features were generated to reduce the overhead, using the data that entered the destination network. Along with this, the "triangle area map" module was applied to extract the geometrical correlations from a pair of two distinct features to increase the accuracy of zero-day attack detection. In an attempt to improve their results from Ref. [123], Tan et al. [124] used Earth Mover's Distance (EMD) to find the dissimilarities between observed traffic and a pre-built normal profile. The network traffic was interpreted into images by feature extraction using MCA and analyzed to detect anomalies using KDDCup99 and ISCX datasets. Using the sample-wise correlation, the accuracy of their results obtained was 99.95% (KDD) and 90.12% (ISCX). However, the study neither revealed the data size nor the effects of varying sample sizes. Moreover, MCA assumed the change to be linear, which was not a realistic approach. Another form of DoS attack in IoT is called a botnet attack, which was explained earlier in Section 2. To prevent botnet attacks against HTTP, Message Queuing Telemetry Transport (MQTT), and DNS, the authors of Ref. [97] developed an IDS, which is an ensemble of DT, NB, and Artificial Neural Network (ANN). Since the correntropy values of benign and malicious vectors were too close, it was decided to use DT, NB, and ANN as they could classify such vectors efficiently. The performance metrics were detection-rate and false-positive rate, for which their proposed ensemble was better than every individual algorithm in that ensemble. For the datasets of University of New South Wales (UNSW) and Network Information Management and Security Group (NIMS), the accuracies achieved were 99.54% and 98.29%, respectively.

Similar to DoS attacks, the MiTM attacks are one of the most frequently occurring attacks in an IoT network. In regard to this, a lot of technical solutions have been proposed for several applications. The authors in Ref. [4] have used LSTM RNN to prevent the impersonation attacks in a smart healthcare scenario, since traditional feedforward neural networks cannot capture the sequence and time-series data, due to their causal property. Moreover, the researchers solved the vanishing gradient issue of RNN algorithm and improved accuracy. At first, the predicted value was calculated based on the dataset log of three months (for a patient who is taking insulin injections). If the predicted and calculated values differed for more than a certain threshold, then by using the combination of DL and gesture recognition, the correct dosage was ensured. However, the detail of the model and analysis was missing in their work.

Similarly, in another scenario to prevent the impersonation attacks, the authors of Ref. [24] utilized Physical Unclonable Function (PUF), which is an inherent characteristic of silicon chips that is unique and can be used as a basis of authentication in RF communication. During the manufacturing phase, every transmitter inherits some unique features called *offset* from an ideal value. The authors have used these offsets as their features to recognize the device, train their system on it, and then detect the accuracy. Using ANN MATLAB toolbox, the performance metrics were calculated. With the help of ML, the simulation results could detect 4,800 nodes transmitters with an accuracy of 99.9% and 10,000 nodes under varying channel conditions, with an accuracy of 99%. The proposed scheme can be used as a stand-alone security feature, or as a part of traditional multi-factor authentication. PUF is inherent and inexpensive and can significantly benefit IoT, wherein each wireless sensor's physical values can be stored in a secure server replacing traditional key-based authentication. However, the authors in their approach have assumed the server storing the PUF values is safe. Aminanto et al. used an unsupervised ensemble of ML algorithms using SVM, ANN, and C4.5 for feature extraction and ANN as the classifier [11]. In their process of deep-feature extraction and selection (D-FES), first, they used SAE to extract the features, then SVM, ANN, and C4.5 were used for feature selection, and finally, ANN was used to classify. The study achieved an accuracy of 99.92% by using AWID dataset, on which an earlier study by Kolias et al. [20] had the worst accuracy for impersonation attack.

According to *Statista* [35], mobile phone users would reach close to three billion by 2020. This increase in usage made mobile phones vulnerable to the malware attack [12, 13, 40, 115, 133, 135]. Azmoodeh et al. [13] believed that OpCodes could be used to differentiate benign-ware and malware. Class-wise Information Gain (CIG) is used for feature selection because the global feature selection causes imperfections, and even reduces system efficiency especially when the dataset is imbalanced. They also claimed that this combination of OpCode and DL for IoT had never been explored. Using Eigenspace and deep convolutional networks algorithms, 99.68% accuracy was achieved, with precision and recall rates of 98.59% and 98.37%, respectively. Similarly, to mitigate malware, Wei et al. [135] extracted the features using the dynamic analysis technique. They used application functional classification to train the classifier for clean and malicious data, while, in the testing phase, kNN was used to divide data into known categories. J48 decision tree and NB were used to perform 10-fold cross-validation. Depending on the performance metric, the study claimed 90% accuracy.

Contrary to dynamic analysis [135], the authors of Ref. [12] used static analysis techniques for feature extraction considering all the API that were not studied previously. Feature selection was made manually based on the most-used features by the previous researchers. They claimed the accuracy of 98.9% with the second biggest malware testbed dataset ever used. As the intrusion techniques were getting sophisticated, the static analysis became invalid, and it was therefore required to use a dynamic scheme [40]. With the static analysis techniques, the attackers adopted deformation technologies, which could bypass the detection while dynamic analysis methods were promising due to its resistance to code transformation techniques. The authors of Ref. [40] proposed a new framework, called EnDroid, based on these issues. The proposed model used "Chi-Square" for feature extraction, five different algorithms (decision tree, linear SVM, extremely randomized trees, random forest, and boosted trees) as an ensemble for base-classification, while LR was used as meta-classifier. For the dataset, a combination of "AbdroZoo" and "Drebin" datasets was utilized so that an accuracy of 98.2% was achieved. Wang et al. argued that most of the existing literature on malware detection was based on static string features, such as permissions and API usage extracted from the apps [133]. However, since malware had become sophisticated, using a single type of static feature might result in a false-negative. In their proposed model, DriodEnsemble, a fusion of string and structural features was utilized to detect Android malware. Using an ensemble of SVM, kNN,

and RF, the model was evaluated against 1,386 benign apps and 1,296 malapps. The study proved to have attained an accuracy of 98.4%, which was better than detection accuracy (95.8%) using only string features, while the accuracy obtained with only structural features was 90.68%.

Anomaly detection is a generic technique where any irregular traffic is flagged as a threat. Several studies [10, 63, 88] have attempted to provide secure IDS using ML algorithms. In this regard, an unsupervised DL technique called Self-Taught Learning (STL) was used by Niyaz et al. [63], and it was based on SAE and SMR. By using NSL-KDD dataset, the comparison was made using 2-class, 5-class, and 23-class classification, and proved 2-class classification to be better than SMR. A multi-class ML-based classification using Mutual Information (MI) was proposed by Ambusaidi et al. [10]. For the linearly dependent variable, Mutual Information Feature Selection (MIFS) with Linear Correlation Coefficient (LLC) was used. For the non-linear dependent variable, the authors used FMIS+MI, made changes to the already existing MIFS algorithm [103], and showed their novelty. For the Linear model (Flexible Linear Correlation Coefficient-based Feature Selection [FLCFS]), the study modified the existing LLC [103] and proposed a new model. An MI can cope with linear as well as non-linear dependents. However, its algorithm can cause redundancy to the classification. Ambusaidi et al. [10] chose "estimator," which relied on estimating the entropies of the given data using average densities from each datum to its k-nearest neighbors. Another reason for this study was that the previous studies had not provided any steps as to how they chose β. The performance was compared using three different datasets of KDDCUP99, NSL-KDD, and Kyoto 2006+, while the metric performance indicators were Accuracy, DR, FPR, and F-measure. Maimo et al. [88] focused on 5G application for anomaly detection based on LSTM. Features extraction was made from network flows using weighted loss function, while feature reduction was made by using DBN and SAE models because of similar structure (where the prediction can be computed using matrix operations followed by the activation function) [88]. After implementing their model using CTU-13 botnet dataset, the authors claimed to have obtained a precision of up to 0.95.

Several studies using ML algorithms as a tool have claimed to reduce cyber-attacks effectively. However, Zhou et al. [146] based their proposal *Deep Feature Embedding Learning* (DFEL) on DL because traditional ML algorithms took extra time to train data. The comparison of their proposal using the datasets of NSL-KDD and UNSW-NB15 confirmed the improvement in recall level of Gaussian Naive Bayes classifier from 80.74% to 98.79%, apart from the running time of SVM significantly reduced from 67.26 seconds to 6.3 seconds. In another similar study [101], the authors claimed that the existing ML algorithms were inefficient for IoT applications and, therefore, a much faster extreme-learning-machine (ELM) could be used instead [101]. Furthermore, they found that the existing security approaches for IoT were centralized and cloud-based, and they, in turn, inherited latency and high power consumption. The proposed IDS for IoT used fog computing for implementation in a distributed fashion in two steps. In the first step, attack detection at fog nodes used an online sequential extreme learning machine (OS-ELM) to identify the attacks in the incoming traffic from the IoT virtual clusters. In the second step, these detected threats were summarized and analyzed at a cloud server. The results of the new algorithm showed better accuracy, False Positive Rate (FRP), and True Positive Rate (TPR) after comparison with the existing NB, ANN, and standard ELM. Furthermore, the experimental results using the Azure cloud also confirmed that the fog-computing-based attack detection was faster than the cloud-computing-based attack detection. However, the study did not compare the results with any existing ML/DL based algorithm used for fog-computing.

4.1.2 Privacy Efforts. Several privacy-preserving ML algorithms have been proposed, as shown in Table 8. Similar to security, privacy is also compromised by an MiTM attack. In this regard,

Table 8. Existing IoT Privacy Solutions Using Machine Learning Algorithms

Ref.	Threat	ToA	Use Case	Algorithm	Dataset	Accuracy
Xiao et al. [137]	MiTM	Spoof detection	WSN	QL, DQ	Private	-
Xiao et al. [139]	MiTM	Spoof detection	MiTMO Landmark	Softmax	Private	-
Aksu et al. [7]	MiTM	Authentication	Wearable devices	best of 20	Private	(Precision) 98.5%
Ma et al. [86]	Data Privacy	Data leakage	Cloud	SGD	-	95%
Zhang et al. [143]	Data Privacy	Inference attack	VANET	LR	NSL-KDD	-
Jia et al. [66]	Data Privacy	Multiple	Distributed Systems	OMPE	Realworld	-
Zhu et al. [147]	Data Privacy	Multiple	Healthcare	SVM	Realworld	94%
Sun et al. [121]	Data Privacy	Multiple	General	HBD, NB, DT	-	-
Feng et al. [39]	Anomaly	Spam	MSN	CNN	Sino Weibo	91.34%

Here, ToA means type of attack.

several studies have used ML algorithms to counter different types of MiTM attacks. For example, the study by Xiao et al. [137] used game theory—-a kind of reinforcement learning, which compared the channel states of the data packets to detect spoofing attacks. The authentication process was formulated as a zero-sum authentication game consisting of the spoofers and the receivers. The threshold was determined by using Nash Equilibrium (NE), implemented over universal software radio peripherals (USPRs), and the performance was then verified via field tests in typical indoor environments.

As an improvement to their work, Xiao et al. [139] applied logistic regression to evaluate the channel model information collected from multiple access points to detect spoofing more accurately. A comparison was made using distributed Frank-Wolfe (dFW)-based and incremental aggregated gradient (IAG)-based authentication to reduce overall communication overhead. IAG-based PHY-layer authentication reduced communication overhead and increased detection accuracy. The results showed improved FAR, DR, and computation costs by using a real-time dataset. In addition to authentication issues, Aksu et al. [7] raised an argument concerning the wearable device, for which the previous schemes only focused on user authentication. However, the device being used should also be authenticated. Such devices could act as MiTMs, which might have similar user authentication details. However, in the background, it might leak all the information to the attacker. Wearables could only connect to the more powerful base device via Bluetooth with authentication and encryption. Since the device name and encryption keys could be compromised easily, it was therefore much more secure to use hardware-based fingerprinting [7]. The proposed framework in Ref. [7] utilized an inter-packet timing-based timing analysis method based on the Bluetooth classic protocol packets. There were four steps in this framework. The first step captured Bluetooth classic packets. The second step extracted the features. In the third step, using probability distributions, the fingerprints were generated. Moreover, as a final step, the stored fingerprints in step three were compared with any new incoming data from wearable devices, to identify any unknown wearable device. By selecting the best algorithm out of 20 from the training results, the study claimed to achieve an accuracy of 98.5%.

Data plays a crucial role in training an ML model. For example, we can use patients' historical data to make a predictive decision for any new patient. However, patients are reluctant to share their data due to obvious privacy concerns. The studies, as shown in Refs [66], [86], and [147], have worked toward solving these issues. In Ref. [147], the researchers proposed a new framework called eDiag, which used non-linear kernel SVM to successfully classify medical information, while preserving user data and service provider's model privacy. Previous studies had used HE techniques, which, according to the study, were not appropriate for online medical prediagnosis. Using their framework, Zhu et al. [147] claimed to have achieved a classification accuracy of

94% without compromising privacy. Similarly, the authors in Ref. [66] classified the privacy issues as *learning-privacy problem* and *model-privacy problem* to protect users' sensitive information and model results, respectively.

Jia et al. [66] argued that the previous work used either gradient-values instead of real-data, or they assumed that the learning model was private, but the learned model was publicly known, or they used complicated encryption procedures. In comparison to all of these studies, Jia et al. [66] proposed a uniform Oblivious Evaluation of Multivariate Polynomial (OMPE) model, which did not contain complicated encryption procedures. Their results proved that the classification data and learned models were protected from several privacy attacks. The research in Ref. [66] focused on model-privacy issues. However, the learning-privacy problem was not discussed. This issue was solved by Ma et al. [86], who argued that encrypting any user-data by the public key was a widely used privacy-preserving technique but at the cost of key management. To preserve the data privacy, Ma et al. [86] proposed a cloud-based DL model that worked with multiple keys to attaining privacy of the user data called Privacy-preserving DL Multiple-keys (PDLM). In their proposed model, a service provider (SP) sent encrypted user data to the cloud, which performs training of the data without knowing the real data. Their evaluation of the PDLM showed that PDLM had successfully preserved privacy with lower efficiency as compared to the conventional non-private schemes.

To improve ML algorithms privacy, Sun et al. [121] proposed an improved version of fully HE that reduced the size and noise of the multiplicative cyphertext by using the re-linearization technique. In their scheme, private hyperplane decision-based classification, private Naive Bayes classification, and private decision tree's comparison were also implemented. In a similar paper, the same authors successfully reduced the user-server iterations to half, without compromising privacy.

Social media platforms like Twitter and Facebook have enriched people's lives at the cost of privacy issues. Several companies used blacklisting techniques to filter benign traffic. However, a survey showed that 90% of the people would fall prey to these attacks before they were blacklisted. To prevent these attacks efficiently, ML algorithms were used. However, these algorithms were inefficient in real-time due to their slower learning rate. In a study, Feng et al. [39] proposed a multistage detection framework using DL, where an initial detection occurred at a mobile terminal whose results were then forwarded to the cloud server for further calculation. By using CNN as a classification algorithm, the authors claimed to achieve approximately 91% utilizing the Sino Weibo dataset. Similarly, the lack of privacy protection mechanisms in a VANET environment was raised by Zhang et al. [143]. In VANET, Vehicle nodes tend to learn collaboratively, raising privacy concerns, where a malicious node can obtain sensitive data by inferring from the observed data. A single node has limited computational and memory resources. The solution was presented by using collaborative IDS with distributed ML algorithms and resolving the privacy issues by proposing the concepts of dynamic differential privacy to protect the privacy of a training dataset.

4.2 Existing Solutions Using Blockchain Technology

Blockchain (BC) is a secure mesh network [15], which is fault-tolerant, transparent, verifiable, and audit-able [25]. The frequently used keywords to describe BC benefits are *decentralized, P2P, transparent, trust-less, immutable*. These attributes make a BC more reliable than an untrusted central client-server model. The smart contract is a computer protocol on BC, which guarantees the execution of a planned event [23]. According to Restuccia et al. [106], the blockchain guarantees data integrity and validity, making it a suitable solution for protection against data tampering in IoT devices.

4.2.1 Security Efforts. Several BC-based solutions for supply-chain, identity management, access management, and IoT were proposed [75]. However, the existing solutions either do not respect the time delay, and cannot be applied to the resource-constrained IoT devices [87]. In contrast to that some studies, like Ref. [126], were only focused on the improvement of time response of an IoT device, rather than their security and privacy. Machado et al. [87] offered data integrity for Cyber-Physical Systems (CPS) by splitting their BC architecture into three levels: IoT, Fog, and Cloud. At the first level, the IoT devices in the same domain created trust in each other using Trustful Space-Time Protocol (TSTP), which is based on Proof-of-Trust (PoT). At the Fog level, Proof-of-Luck (PoL) was used to create fault-tolerant IoT data, which produces a cryptographic digest for a data audit. The data generated from the first level was hashed using SHA-256 and saved temporarily. After the acknowledgment and consensus were reached, the data was permanently stored at the third level of cloud, which is a public ledger. Other than data integrity, the study also offered key management using time synchronization and the location of the node. Heuristic Environmental Consideration Over Positioning System (HECOPS) was used to estimate the node's location via multi-lateration, and TSTP provided clock synchronization. The paper proposed to use multiple consensuses, such as PoT and PoL, but it did not cater to any user privacy issue. Another paper [83] provided data integrity with the idea of securing data collected from the drone using public BC. DroneChain presented had four modules; *drones*, *control system*, *cloud server*, and *a BC network*. Drones were controlled by the control system, and the data was encrypted and stored using the cloud server on a decentralized BC. The resultant system was trusted and accountable, offered instant data integrity, and had a resilient backend. However, the study used PoW, which was not the best choice for a real-time IoT application like drones. In addition, the work did not offer data provenance and user/data security.

DoS attacks are one of the frequently executing attacks due to their comparatively straightforward implementation and the ever-growing number of insecure digital devices. Due to cheap IoT technologies, hackers can easily control multiple IoT devices to launch an attack. According to Ref. [128], the SDN top layer is prone to brute force attacks. Since SDN is controlled by software, it can be targeted by injecting malicious applications and also gives rise to the DoS/DDoS attacks. The earlier methods to prevent DDoS are not compatible with a light-weight multi-standard IoT environment. Other than that, SDN can suffer flooding attacks, saturation attacks, and MiTM attacks due to lack of authentication in the plain-text TCP channel. Tselios et al. [128] argued that BC offered a better solution to protect IoT devices from security attacks and enforced trust between multi-vendor devices, as it was decentralized, fault-tolerant, and tamper-proof. These valuable BC properties make it resistant to data tampering and flooding attacks. However, all of the solutions mentioned above were theoretical ideas as no practical implementation was done. In another paper, Sharma et al. [113] improved the security vulnerability in SDN by proposing a distributed SDN architecture for IoT using BC called DistBlockNet. The BC was used to verify, validate, and download the latest flow rule table for the IoT forwarding devices. The proposed DistBlockNet model was compared with the existing solutions, and the results were better in terms of real-time security threat detection and overhead usage.

In another study, the researchers highlighted a MiTM security gap in a smart-grid, where any malicious actor could modify user data sent over the Internet [44]. Secondly, the customers could not audit their costly utility bills, because the current smart-grid was unpredictable, and it did not provide any early warnings to the customer indicating higher energy usage. To avoid the above issues, this study proposed to use cryptographic data transmission using public and private keys for the user ID as well as the smart contract, which was placed on a BC. This technique ensured an immutable, secure, and transparent smart-grid system. However, PoW could be extremely expensive and resource exhausting.

The study in Ref. [55] argued that the existing logistics systems were neither transparent nor credible to trace. The existing systems were centralized, relied on multiple TTPs, and focused on a single transporter. Hasan et al. [55] proposed a proof of delivery system using BC technique. In their transporter system, the nodes were *seller*, *buyer*, *courier services*, *arbitrator*, and *Smart Contract Attestation Authority* (SCAA). The initial agreement was a smart contract that was placed on Inter-Planary File System (IPFS) and was executed once all the parties agreed. The item was transported between several transporters as per the smart contract (a maximum of three in this paper), which was created every time for the next transporter. Finally, once the buyer has verified and collected the item, the payment is released to the seller. In the case of any rejection (i.e., transaction failure), the *arbitrator* takes over, settles the dispute, and redistributes the amount based on the negotiated agreement. This proposed physical-asset-delivery system has inherent BC security against MiTM and DoS attacks. However, the authors have not paid any particular attention to user ID management and data privacy. The study by Gupta et al. [54] was a simulation done in OMENT++ on one application scenario where the authors claimed to have tackled Sybil attacks as well as the replay attacks in an IoT network. First of all, they introduced a new layered architecture, which had two more layers in the underlying IoT architecture. They explained their algorithm, idea, and work by comparison in terms of metrics of *Transactions added to the BC per second (Ftx)*, *Blocks added to the BC per second (Fblk)*, and *Memory space utilized (Mmempool)*.

IDS is one of the widely used monitoring devices to detect anomaly traffic behavior. In a study by Golomb et al. [49], the authors argued that the current anomaly IDS were not efficient since the training phase considered only benign traffic. An adversary could exploit this vulnerability by injecting malicious data, which might be regarded as benign. Secondly, the trained model might not be as efficient, since it might be missing some IoT device traffic, which was only event-driven by, for example, a fire alarm. Both of the issues were solved by using a Collaborative IoT Anomaly (CIoTA) Detection using BC technique, where all IoT devices of the same type were trained simultaneously. Since a large number of IoT devices were being trained based on their local data traffic, the chances of an adversarial attack were minimum. Each device would generate a locally trained model, which would be collaboratively merged into a globally trained model by using BC technique. The study successfully implemented CIoTA and proved its benefits for eliminating the adversarial attacks. However, the separate block generated for each IoT model would increase the amount of data.

Along with the research on frequently researched security threats such as Data integrity, MiTM, and DoS, several studies have focused on providing solutions to multiple attacks. Sharma et al. in Ref. [112] presented an affordable, secure, and always accessible BC technique for distributed cloud architecture. The combination of SDN and BC implemented the security of the fog nodes. The study brought the resource extensive tasks closer to the edge of an IoT network, which not only ensured better security but also improved end-to-end transmission delay. The authors further claimed that the model was adaptive based on the encountered threats and attacks, and reduced administrative workload. The main focus of this paper was to provide an architecture based on BC-cloud in fog computing, which was scalable, secure, resilient, and fast. The comparison was made in terms of throughput, response time, and false alarm rate. However, there was no consideration to the data privacy, user ID management, or the key management. Similarly, Sharma et al. in Ref. [114] claimed that the existing Distributed Mobile Management (DMM) lacked robustness against the security threats due to its centralized architecture. Their proposed scheme based on the BC showed improved latency, delay, and energy consumption, without affecting the existing network layout. However, the study used PoW consensus, which is energy-hungry and offered no user privacy.

Table 9. Taxonomy of Existing IoT Security Solutions Using Blockchain Techniques

Ref.	Threat	Use Case	BC used	BC type	Consensus	Security	Weakness
Machado et al. [87]	Data Integrity	Cyber Physical System	Ethereum	Public	PoT + PoL	D/K	Did not address *U*
Liang et al. [83]	Data Integrity	Drone	-	Public	PoW	D/K	(i) PoW is inefficient for real-time applications (ii) Public BC is insecure
Tselios et al. [128]	DoS	SDN	NG	Public	-	None	*U/D/K* not addressed
Sharma et al. [113]	DoS	SDN	Bitcoin	Public	PoW	None	Lack of data integrity & *U*
Gao et al. [44]	MiTM	SmartGrid	-	Private	PoW	U/D/K	Encryption techniques are complex and slower
Hasan et al. [55]	MiTM	logistics	Ethereum	Private	PoW	K	Did not address *U* & *D*; Overall less secure
Gupta et al. [54]	MiTM	IoT	Bitcoin	Public	Private	K	Only simulation is done for basic security
Golomg et al. [49]	Anomaly	Network	Private	Public	Private	D/K	Block per IoT model will increase the data.
Sharma et al. [112]	Multiple	Fog-SDN	Ethereum	Public	Proof-of-Service	None	No *U* or *D* is offered
Sharma et al. [114]	Multiple	5G	Multiple	Both	Multiple	None	PoW is costly, plus *U/D/K* not addressed

Here, U, D, and K mean *User security*, *Data security*, and *Key management*, respectively.

All of the above solutions are mentioned in Table 9, where most of the researchers have focused on using PoW as a consensus algorithm, which is not suitable for a real-time IoT application. Moreover, most of them have not considered user anonymity and data integrity.

4.2.2 Privacy Efforts. Privacy is a complicated issue in a BC that can be accomplished, but at the cost of throughput and speed [25]. A hacker can identify the patterns of a permissionless BC since all of the transactions happen in public and make an informed decision about the source. BC-based privacy-preserving was proposed by several researchers to solve this issue [5, 37, 53, 69, 79, 85, 105, 131, 145].

Wang et al. proposed a BC-based model, tackling the MiTM attack issues in a crowdsensing application [131]. The user privacy was implemented by using node cooperation method, in which the server released the sensing task as well as its price, which was pre-paid on the BC. The users would perform the sensing task and upload the sensing data, and finally, the user was paid as per their achievements. To achieve user-data privacy, the authors proposed k-anonymity, in which the sensing task was not given to an individual, but a group and the sensed data gathered was also in the form of a group, which preserved privacy of a single-user. The announcement VANET is something in which the users (nodes) shared some information that might benefit other users in the network. According to the researchers of CreditCoin [79], the current VANET system had a lack of privacy as well as motivation for the users to share any data. CreditCoin was proposed that offered decentralization, trust, and motivation by paying the user their incentives. The shared information was immutable, so the source did not fake any news either, benefiting the whole VANET community from it. For example, the information might be "a traffic accident on ABC road going toward XYZ". Another VANET application was proposed by Lu et al. in Ref. [85], where the authors added privacy to the users in the existing Bitcoin platform using the lexicographic Merkle tree. Furthermore, the forgery was controlled by adding a reputation weight to every vehicle in the network. However, the study used PoW as their consensus protocol, which is very costly and can create traffic bottlenecks in a resource constraint VANET application.

First, of its nature, Zhou et al. [145] claimed to design the BC-based IoT system where the servers helped users to process encrypted data without learning from the data. HE was used to secure the data in a private BC using PBFT consensus. The authors in Ref. [105] argued that although the BCs were immutable and tamper-proof, once a block was executed, they did not cater confidentiality and privacy of the data as anyone could see the plain-text. When such a BC was integrated with IoT, it was more vulnerable due to a massive influx of data. Rahulamathavan et al. focused on these issues by proposing a privacy-preserving BC architecture for IoT applications based on the Attribute-based Encryption (ABE) [105].

The previous studies offered the solution by using symmetric encryption like AES, which meant that the key must be shared with the data to enable the miners of the BC to verify the content and update the BC. However, such a technique could not guarantee privacy. ABE used single encryption to keep data private and safe. In a scenario of a hospital, the main server could encrypt data before transmitting the attributes, such as DOCTOR or NURSE, which could only be read by the concerned node by using the same attributes and decrypting them. The BC architecture could secure data manipulation since multiple nodes verified a single transaction. After the approval, the data was stored and could not be tampered. Lastly, there was no central control, making all of the transactions transparent and fair. However, the cluster head could read the data, which might be exploited by an attack.

Fan et al., working in the 5G network application, argued that the work on access control of encrypted data still needed to be explored [37]. Despite several advantages of ABE, if a user wanted to change his policy, the attribute revocation and re-encryption took much time. Additionally, the owners did not control their public data, and the trust was delegated to the third parties. Centralized systems were fault-prone, and could cause traffic choking. Fan et al. used BC to solve these issues by using encrypted cloud storage for the provision of privacy-preserving and data-sharing systems, which was tamper-resistant, fully controlled by the user, and always accessible to anyone on request [37]. However, their proposal had several drawbacks; for example, the miners could share the information without user consent. Moreover, the BC proposed is public, which means anyone could access it.

Aitzhan et al. [5] addressed the issues of transaction security and privacy by using multi-signatures. Since the traditional systems were insecure, unreliable, and publicly accessible, the messages were sent in an encrypted form that offered privacy and security in communication. User anonymity was ensured by using the public key and private key. Similarly, another concept of multi-signatures was mentioned by Guo et al. [53]. The authors found that the current Electronic Health Record (EHR) system was centralized with no user privacy or control over it. Health records are critical documents as they have a personal medical history. The user should be in control of them, but they should be unforgeable as well. In previous studies, Attribute-Based Signatures (ABS) enabled trust between the two parties; however, it was unreliable and restricted to a single signature. Encashing the ABS advantages, Guo et al. presented an ABS with multiple access (MA-ABS), which guaranteed privacy with access control to the user, and confidence of real information to the verifier [53]. Moreover, using BC for maintenance of data reinforced immutation, unforgeability, and decentralization. Privacy-preserving was achieved by using MA-ABS and collusion attacks were avoided by using pseudorandom function seed. The study also proposed Key management by using KeyGen.

In a similar attempt, Ref. [69] offered a new consortium BC called P2P Electricity Trading system with COnsortium blockchaiN (PETCON), which was based on the Bitcoin platform using PoW for the PHEV to trade the surplus electricity between them. The existing P2P was a single point of failure, and it was expensive and untrustworthy. Kang et al. [70] improved upon the privacy of a vehicular data in the existing P2P data sharing networks. Due to the resource

constraints in a vehicular system, the data was forwarded to the edge computers for powerful computation. The data shared was vulnerable, due to which the researchers in this study used consortium BC, where only the selected nodes could perform the audit and verification. They also introduced the use of smart-contracts, which ensured user-authenticity and secure data-sharing, and improved data-credibility. The consortium model reserved the energy as it selected a lesser number of nodes for data maintenance. Vehicle-ID authentication was done by digital signatures using public/private keys, while *Elliptic curve digital signature algorithm* provided key-management. The authors also touched upon data privacy management by storing the raw data using the proof-of-storage.

4.3 Existing Solutions Using Machine Learning and Blockchain

In this section, we look at the existing security and privacy solutions for IoT with the integration of ML algorithms and BC techniques, as depicted in Tables 11 and 12.

4.3.1 Security Solutions. Agrawal et al. claimed to eliminate spoofing attacks with the combination of ML algorithms and BC techniques [3]. By securing the user-device communication, the user in a valid IoT-zone is continuously monitored, and the communication logs are saved on the BC. The records are immutable and can be verified for any suspicious activities. The existing user authentication techniques include one-time-password (OTP) or security questions, which are limited to single authentication. By using Hyperledger as a BC platform, the authors resolved this issue by considering continuous security using IoT-zone identification, IoT-token generation, and token validation. However, the study considered IoT-hub as a center of communication, which voided the concept of decentralization. There was no user or data privacy in concern, and the dataset was too small for a DL model.

The open nature of Android poses new security challenges and attacks. Gu et al. [52] illuminated that Android-based systems were highly targeted by malware, trojans, and ransomware with evolving nature when studied overtime [144]. The existing schemes, which can be classified as either static-based analysis or dynamic-based analysis, had certain drawbacks such as high computation time costs and types of code obfuscations such as variable encoding and encryption [61]. Gu et al. proposed a new multi-feature detection model (MFM) of Android-based devices, where they utilized a fact-base of malicious codes by using Consortium BC for Malware Detection and Evidence Extraction (CB-MDEE) in mobile devices. Compared with the previous algorithms, CD-MDEE achieved higher accuracy with lower processing time.

Using the Exonum BC platform and DNN ML algorithms, the proposed architecture leverage upon BC's properties to send and sell their data *as and when* required giving optimum access control to their health data [90]. As the data in the storage would be encrypted, the compromise of the storage would not lead to data leakage. The proposed scheme utilizes hash functions and public-key signatures for encrypting user data to guarantee authorization and validity. The paper, however, lacks the in-depth comparison with other schemes, other than being just a theoretical framework.

4.3.2 Privacy Solutions. Many companies rely on big datasets to optimize their target audience and enhance their profits, but such data contain sensitive personal information, such as political preferences, which can be exploited by interested entities. It is, therefore, crucial to preserve the privacy of such users, and if required, compensate them for their contributions. Moreover, certain domains have an abundance of data, which can be beneficial for research and development, but the data cannot be shared with third parties. Furthermore, the same data can be manipulated and raise doubts on its integrity. To improve upon the above architecture, several studies have been proposed [36, 91, 93, 94, 116].

Mendis et al. [92] proposed fully autonomous individual contributors working in a decentralized fashion without disturbing the functionality and overall efficiency, which they later on improved in their work in Ref. [93]. Their comparison against federated learning using the MNIST dataset for CNN model generated more than 94% accuracy in each scenario. The smart contracts incentivizing the computing contributors executed the peer-to-peer transactions. However, in their study [93], the execution time with encryption increased 100%. Moreover, the architecture was based on the ethereum BC having a block-time of 12 seconds, and, hence, it might not be feasible for a real-time IoT application, for example, video streaming.

DeepChain proposed BC based value-driven, incentives mechanism to solve security issues [136]. DeepChain guarantees data privacy and audit-ability for the model training process. Confidentiality is employed using the Threshold Paillier algorithm that provides an additive homomorphic property. Using CNN algorithms and MNIST dataset, DeepChain proved that the more parties that participated in collaborative training, the higher the training accuracy was.

ML classifiers require datasets to train. These datasets are collected from different entities who are usually reluctant to share their data due to several privacy concerns such as data leakage, data integrity, and ownership. The users do not know how and when their data may be used. To preserve these privacy issues, Shen et al. [116] proposed a fusion of machine learning with blockchain. A privacy-preserving SVM-based classifier was used to train the encrypted data collected from IoT users, while the BC platform provided data sharing among multiple data providers. However, the solution used encryption techniques to preserve privacy, which is not suitable for a resource constraint IoT device. The use of the BC platform is also not explained in detail.

In yet another study, an attempt to create tamper-proof DNN models is done with the help of BC [46]. Using the BC properties like *transitive hash*, *cryptographic encryption*, and *decentralized nature*, an architecture named *DeepRing* is proposed. A shared common ledger stored the state of the model. Ouroboros block stored all blocks' hashes, which was used to track the compromised block in case of any tampering attack. Since the querent encrypted the query with its public key, and the output was only encrypted using the public key of the querent, no one else could access the model results. Focusing on the adversarial attacks on network parameters, the authors compared DNN architecture with DeepRing architecture. The DNN architecture without BC using CIFAR-10, MNIST, and Tiny ImageNet datasets dropped by their accuracy by 20.71%, 47%, and 34%, respectively. However, the DNN with BC suffered 0% accuracy loss.

Similar work is done in the latest research by Fadaeddini et al. [36], who proposed a framework where the privacy of data-owners was preserved by training the shared model on their data locally. After the learning is completed, the data-owners only shared the learned parameters of the model. The study demonstrated *self-driving cars* application scenario, which used the Stellar BC platform for the decentralized deep learning infrastructure. The contributors are paid for their work as they helped in improving the accuracy of self-driving cars. The learned model is saved on a distributed file system known as IPFS (Inter-Planary File System), which is resistant to DDoS attacks. The framework also controls the authenticity of computing partners to avoid any malicious activities. Although the work is novel and ticks all the privacy issues (i.e., user privacy, data privacy, and key management), there is a lack of comparative analysis that can prove that their work is better than the traditional framework.

5 RESEARCH CHALLENGES

5.1 Challenge to Machine Learning Algorithms in IoT

ML algorithms are utilized for analysis after being trained on a large number of datasets to adapt to the desired output dynamically. These models may be used, for example, in navigating a robot or for speech recognition, where human expertise either does not exist or cannot be used. ML

algorithms have also been utilized very efficiently to analyze threats against several cybersecurity domains. Although ML algorithms perform well in many areas, they have some limitations in the IoT environment:

- **Scalability and Complexity:** In recent studies, several ML algorithms have effectively reduced the cyber attacks. However, ML algorithms are not an ideal pick for IoT applications due to its limitations. Diro et al. claimed that the traditional ML algorithms were limited in scalability, feature extraction, and accuracy [30]. Whereas, Moustafa et al. [97] argued that ML algorithms could not solve many problems, primarily when it was implemented in a complex resource-constrained IoT environment. Another work done by Abeshu et al. [1] proved that the traditional ML algorithms were less scalable and less accurate in a vast distributed network such as IoT. After comparing classical ML algorithms with DL methods, several studies learned that most DL techniques used pre-training for feature extraction. DL not only saved administrative time but also reduced feature dimensionality by reducing redundancy [63, 71, 81, 109, 140].
- **Latency:** As a solution to the above issues, some authors, for example, Xiao et al. [139] proposed to use ensemble ML algorithms. The ensemble algorithm proved to be performing better than each ML algorithm individually, but it was computationally expensive. As an alternative to classical ML, most of the studies pointed out that DL is a better choice for IoT. In another study, the authors proposed DFEL [146]. They utilized the DL-based model because the traditional ML algorithms increased training time in Big Data scenarios. Using the datasets of NSL-KDD and UNSW-NB15, they claimed to have improved in the recall of Gaussian Naive Bayes classifier from 80.74% to 98.79%. Moreover, their method significantly reduced the running time of SVM from 67.26 seconds to 6.3 seconds. The improvement in recall-rate and running time perfectly suits an IoT application.
- **Compatibility:** Although the above solutions have performed better, we believe that these DL-based techniques are application-specific. In such cases, a model trained for solving one problem may not be able to perform well for another problem in the similar domain [59].
- **Vulnerability:** One of the critical challenges to the ML/DL techniques in IoT is to secure themselves from any security or privacy attacks. Adversarial attacks against machine learning models may degrade system performance, as such attacks significantly reduce the output accuracy [82]. The attack severity is proportional to the amount of information available to an adversary about the system [22], which is very difficult to counter. As depicted in Figure 5, an adversary can attack ML models at different levels, for example, tampering the input parameters. Goel et al. [46] highlighted that much work is done to counter input level attacks [2, 6, 47, 50, 51]; however, the research focus on adversarial attacks on network parameters is very less. Some of these attacks can be proven deadly, for example, in a healthcare application where an ML algorithm is used to analyze the amount of insulin provided by a patient. If an adversary can inject malicious code and alter the ML algorithm's input, the amount of insulin may be increased and cause death to the patient.

  Regarding the above issues, we believe that the ML algorithms for IoT need to be optimized for scalability, speed, compatibility, and security, and privacy. We think that privacy-preserving ML algorithms, such as differential privacy and light-weight HE, should be explored to overcome the discussed challenges.

5.2 Challenges to Blockchain in IoT

- **Latency and speed:** Although the BC technology was introduced a decade ago, its real benefits were realized only recently. In recent studies, many efforts have been made to utilize

Table 10. Overview of Existing IoT Privacy Solutions Using Blockchain Techniques

Ref.	Threat	Use Case	BC used	BC type	Consensus	Privacy	Weakness
Wang et al. [131]	MiTM	Crowdsensing	Bitcoin	Private	PoW	U/D	Prone to collusion attacks
Li et al. [79]	MiTM	Vanet	Private	Private	Private	U/D/K	Poor key management
Lu et al. [85]	Data Privacy	VANET	Bitcoin	Private	PoW	U/D/K	PoW is slow & not ideal for real-time scenario
Zhou et al. [145]	Data Privacy	IoT	Ehtereum	Private	PBFT	U/D	Block time not suitable for real-time IoT
Rahulamathavan et al. [105]	Data Privacy	IoT	Bitcoin	Public	PoW	D/K	Unsuitable for real-time IoT as block time is 10 m
Fan et al. [37]	Data Privacy	5G	Private	Public	DPos	U/D/K	Miners can share data & store data, BC is public
Aitzhan et al. [5]	Data Privacy	Smartgrid	PriWatt	Public	PoC	U	Did not address *D* and *K*
Guo et al. [53]	Data Privacy	Healthcare	Private	Public	-	U/D/K	No BC model or consensus technique mentioned
Kang et al. [69]	Data Privacy	PHEV	PETCON	Consortium	PoW	K	Did not address *U* or *D*

Here U, D, and K mean *User security*, *Data security*, and *Key management*, respectively.

Table 11. Overview of Existing IoT Security Solutions Using Machine Learning Algorithms and Blockchain Techniques

Ref.	Attacks	Use Case	Algo	Dataset	Metric	BC used	BC type	Consensus	Privacy
Agrawal et al. [3]	MiTM	IoT	VMM+ LST	Private	Accuracy	Hyperledger	Private	PBFT	K
Gu et al. [52]	Malware	Android	MFM	Drebin	FPR, DR, Acc	Private	Consortium	-	none
Mamoshina et al. [90]	Access Control	Healthcare	DNN	-	-	Exonum	Private	BFT	U/D/K

Here, K stands for *Key management*.

Table 12. Summary of Existing IoT Privacy Solutions Using Machine Learning Algorithms and Blockchain Techniques

Ref.	Attacks	Use Case	Algo	Dataset	Metric	BC used	BC type	Consensus	Privacy
Mendis et al. [92]	Data Leakage	General IoT	CNN	Private	Accuracy	Ethereum	Private	PoS	D
Mendis et al. [93]	Data Leakage	SDN	CNN	MNIST	Accuracy	Ethereum	Private	PoS	U/D/K
Weng et al. [136]	Data Privacy	General	CNN	MNIST	Accuracy	Corda	Private	BAP*	U/D/K
Shen et al. [116]	Data Privacy	Smart Cities	SVM	BCWD+HDD	Accuracy	NG	NG	PoW	U/D/K
Goel et al. [46]	Data Tampering	Computer Vision	DNN	MNIST/CIFAR-10	Accuracy	Private	Public	-	U/D/K
Fadaeddini et al. [36]	Data Privacy	Self-driving Cars	-	-	-	Stellar	Public	SCP†	U/D/K

Here, U, D, and K mean *User security*, *Data security*, and *Key management*, respectively.
*Byzantine agreement protocol.
†Stellar Consensus Protocol.

BC in several applications, such as logistics, food, smart grid, VANET, 5G, healthcare, and crowdsensing. However, the existing solutions do not respect the latency issues of BC, and cannot be applied to the resource-constrained IoT devices [31, 87]. The most widely used BC consensus is PoW, as depicted in Table 10. PoW is slow (limited to 7 transactions per second compared to an average of 2000 transactions per second for the Visa credit network) and requires a lot of energy [16, 23, 25].

—**Computation, processing, and data storage:** There is a substantial cost of computation, power, and memory involved in maintaining a BC across a vast network of peers [16, 119]. According to Song et al., in May 2018, the Bitcoin ledger size had surpassed 196 GB.

These limitations suggest poor scaling and transaction speed for an IoT device. Although an alternative was to offload their computation tasks onto a central server—cloud, or a semi-decentralized server—fog, this, however, adds network latencies [107, 119].

— **Compatibility and Standardization:** Like any emerging technology, one of the BC challenges is its standardization for which the laws need to be reformed [98]. Cybersecurity is a difficult challenge, and it would be naive to think that we all will see a security and privacy standard that can eliminate all risks of cyber-attack against IoT devices anytime soon. Even so, a security standard can ensure that devices meet "reasonable" standards for security and privacy. There are a number of fundamental security and privacy capabilities that should be included in any IoT device.

— **Vulnerability:** Although the BC is non-repudiable, trustless, decentralized, and tamper-proof, a blockchain-based system is only as secure as the system's access point. In a public BC-based system, anyone can access and view the data contents. While the private blockchain is one of the solutions to the above problem, it raises other issues such as trusted third party, centralized-control, and access-control legislation. In general, the blockchain-enabled IoT solutions must meet the security and privacy requirements such as (i) the data must be stored securely by satisfying the confidentiality and integrity requirements; (ii) data must be securely transmitted; (iii) data must be shared transparently, securely, and in an accountable fashion; (iv) the properties of authenticity and non-reputation must be preserved; (v) the selective disclosure property must be satisfied by the data-sharing platform; and (vi) the explicit consent of data sharing must be taken by the involved parties [41].

5.3 Challenges to ML & BC in IoT

We believe that a single technology or a tool, like BC or ML, will not suffice in providing optimum security and privacy for IoT networks. Therefore, it is a dire need of time for the research community to explore the provision of IoT security and privacy with the merger of BC and ML, that has the following challenges:

— **Storage:** As discussed in Section 4, ML algorithms perform better with larger datasets [1, 30]. However, the increase of data in BC platforms will degrade its performance [119]. It is an open research issue to find a balance, which would be ideal for IoT applications.

— **Latency challenges:** Depending upon the scenario, an IoT network may generate a considerable amount of data requiring more time for training and computation, which may potentially increase the overall performance (i.e., latency) of traditional ML models [31, 87].

— **Scalability:** ML and BC have scalability challenges, in terms of both the processing and communication costs. Many ML algorithms impose additional processing and communication costs with the increase of data that is imminent for most IoT networks. Similarly, the BC performs poorly as the number of users and networking nodes increases [29, 111]. On average, an Ethereum BC performs 12 transactions per second, which is unacceptable in traditional IoT applications, where millions of transactions are happening every second [110].

— **Vulnerability:** Although the combination of ML and BC can tremendously increase security and privacy, there are a few challenges as well. The increasing number of threats, including malware and malicious code, increases the challenge of identifying, detecting, and preventing them in real-time IoT networks. The training phase of ML takes longer, and while it is possible to detect malicious traffic, this is only possible with a trained model [82]. Blockchain, on the other side, can guarantee data immutability and can identify their transformations. However, the issue is with the data that is corrupted before entering the

blockchain. Additionally, the malfunctioning of sensors and actuators from the start cannot be detected until that particular device has been tested [107]. Besides the above issues, public BC is prone to privacy evasion techniques as the stored data is publicly accessible and available to all readers. Using private BC is one of the solutions to these challenges; however, this would limit access to a large amount of data required for ML to perform efficiently [110].

The IoT devices can generate a massive amount of data, which should be typically processed in real time. Since the demand for IoT-based BC is different, there is much research going on to bring a new BC that is compatible with IoT. However, the most important limitations on BC are ledger storage and transaction per second (TPS). Although in the latest BCs, such as Hyperledger Fabric where TPS is down to milliseconds, a lot still needs to be done for a BC to work smoothly in the IoT environment. Similarly, in the context of the secure BC model of IoT, the security needs to be built-in, with validity checks, authentication, and data verification, and all the data needs to be privacy-preserved at all levels. We need a secure, safe, and privacy-preserved IoT framework.

6 CONCLUSION AND FUTURE WORK

In this paper, we have reviewed the latest threats to IoT and categorized them into security and privacy. Their effects, type of attacks, the layer of impact, and solutions have been briefly mentioned. We have then comprehensively presented the latest existing literature survey on IoT security and privacy using ML algorithms as well as BC technologies and highlighted their gaps. This paper has presented the current solutions to IoT security and privacy by utilizing ML algorithms, BC techniques, and the integration of both. To better understand the security and privacy issues in an ML, we have also attempted to present an ML threat model for IoT based on the previous studies. Finally, We discuss a few research challenges to ML algorithms in IoT, BC techniques in IoT, and the challenges to the combination of ML and BC in IoT.

The generation, storage, analysis, and communication of data are fundamental to the IoT ecosystem. A holistic approach is in demand, where a vulnerability-free system needs to be built, through measures such as adherence to best practices and continual testing. The system should be able to learn and adapt to the latest trends in threats (zero-day attacks) since malicious activities are dynamic. In this regard, ML/DL can be extremely beneficial in analyzing the traffic. At the same time, the BC can serve as a basis to keep a ledger of logs and communication in an IoT environment. Since this data is immutable, it can be used confidently in the court of law as a piece of evidence.

Among the studies conducted on IoT security and privacy, most of them focused on providing security or privacy. We believe that for a system to be secure, both security and privacy are equally important. Moreover, data privacy is the most critical factor, which can only be valid when considered end-to-end. The current systems lack the integrity of datasets that are used to train a model. Any adversary can tamper these datasets to obtain their desired results.

Currently, the integration of ML algorithms with BC techniques to achieve IoT security and privacy is a relatively new area, which requires further exploration. However, some of the research questions are: (i) Can we use BC to eliminate DDoS attacks in an IoT network by integrating it with ML algorithms? (ii) Can the resource-constrained IoT device leverage upon BC's inherited encryption to perform in real time? (iii) Can BC introduce trust in traditional collaborative ML-based IoT Intrusion Detection Systems? Moreover, several organizations, both public and private, rely on the data generated by IoT devices. How can we trust the data, whether *in motion*, or *at rest*? This question becomes more difficult to answer in a centralized cloud-based IoT architecture. We can extract meaningful data from privacy-preserving ML algorithms, whereas BC can offer security

and trust. In the future, we aim to design and develop a privacy-preserving IoT framework, which will offer privacy-preserving data sharing and privacy-preserving data analysis.

REFERENCES

[1] Abebe Abeshu and Naveen Chilamkurti. 2018. Deep learning: The frontier for distributed attack detection in fog-to-things computing. *IEEE Communications Magazine* 56, 2 (2018), 169–175.

[2] Akshay Agarwal, Richa Singh, Mayank Vatsa, and Nalini Ratha. 2018. Are image-agnostic universal adversarial perturbations for face recognition difficult to detect? In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. 1–7.

[3] Rahul Agrawal, Pratik Verma, Rahul Sonanis, Umang Goel, Aloknath De, Sai Anirudh Kondaveeti, and Suman Shekhar. 2018. Continuous security in IoT using blockchain. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 6423–6427.

[4] Usman Ahmad, Hong Song, Awais Bilal, Shahzad Saleem, and Asad Ullah. 2018. Securing Insulin Pump System Using Deep Learning and Gesture Recognition. In *Proceedings of the 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications and 12th IEEE International Conference on Big Data Science and Engineering, Trustcom/BigDataSE 2018* (2018).

[5] Nurzhan Zhumabekuly Aitzhan and Davor Svetinovic. 2018. Security and Privacy in Decentralized Energy Trading Through Multi-Signatures, Blockchain and Anonymous Messaging Streams. *IEEE Transactions on Dependable and Secure Computing* 15, 5 (2018).

[6] Naveed Akhtar and Ajmal Mian. 2018. Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey. arxiv:cs.CV/1801.00553

[7] Hidayet Aksu, A. Selcuk Uluagac, and Elizabeth Bentley. 2018. Identification of Wearable Devices with Bluetooth. *IEEE Transactions on Sustainable Computing* (2018), 1–1.

[8] Mohammad Al-Rubaie and J. Morris Chang. 2019. Privacy-Preserving Machine Learning: Threats and Solutions. *IEEE Security and Privacy* 17, 2 (2019), 49–58.

[9] Muhammad Salek Ali, Massimo Vecchio, Miguel Pincheira, Koustabh Dolui, Fabio Antonelli, and Mubashir Husain Rehmani. 2019. Applications of Blockchains in the Internet of Things: A Comprehensive Survey. *IEEE Communications Surveys and Tutorials* 21, 2 (2019), 1676–1717.

[10] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan. 2016. Building an Intrusion Detection System Using a Filter-Based Feature Selection Algorithm. *IEEE Transactions on Computers* 65, 10 (Oct. 2016), 2986–2998.

[11] Muhamad Erza Aminanto, Rakyong Choi, Harry Chandra Tanuwidjaja, Paul D. Yoo, and Kwangjo Kim. 2017. Deep abstraction and weighted feature selection for Wi-Fi impersonation detection. *IEEE Transactions on Information Forensics and Security* 13, 3 (2017), 621–636.

[12] Simone Aonzo, Alessio Merlo, Mauro Migliardi, Luca Oneto, and Francesco Palmieri. 2017. Low-Resource Footprint, Data-Driven Malware Detection on Android. *IEEE Transactions on Sustainable Computing* 3782 (2017).

[13] Amin Azmoodeh, Ali Dehghantanha, and Kim-Kwang Raymond Choo. 2018. Robust Malware Detection for Internet of (Battlefield) Things Devices Using Deep Eigenspace Learning. *IEEE Transactions on Sustainable Computing* 3782, c (2018), 1–1. http://ieeexplore.ieee.org/document/8302863/

[14] Mandrita Banerjee, Junghee Lee, and Kim Kwang Raymond Choo. 2018. A blockchain future for Internet of Things security: A position paper. *Digital Communications and Networks* 4, 3 (2018).

[15] R. Baxter, N. Hastings, A. Law, and E. J. Glass. 2008. *5 Future Uses of Blockchain*. Vol. 39. Retrieved from https://www.thestreet.com/technology/cybersecurity/five-future-uses-for-blockchain-14589274.

[16] S. Biswas, K. Sharif, F. Li, B. Nour, and Y. Wang. 2018. A Scalable Blockchain Framework for Secure Transactions in IoT. *IEEE Internet of Things Journal* (2018).

[17] I. Brass, L. Tanczer, M. Carr, M. Elsden, and J. Blackstock. 2018. Standardising a moving target: The development and evolution of IoT security standards. In *Living in the Internet of Things: Cybersecurity of the IoT - 2018*. 1–9.

[18] Magda Brewczyńska, Suzanne Dunn, and Avihai Elijahu. 2019. *Data Privacy Laws Response to Ransomware Attacks: A Multi-jurisdictional Analysis*. Springer, 281–305. https://@doi.org/10.1007/978-94-6265-279-8_15

[19] I. Butun, P. Österberg, and H. Song. 2020. Security of the Internet of Things: Vulnerabilities, Attacks, and Countermeasures. *IEEE Communications Surveys Tutorials* 22, 1 (2020), 616–644.

[20] C. Kolias, G. Kambourakis, A. Stavrou and S. Gritzalis. 2016. Intrusion Detection in 802.11 Networks: Empirical Evaluation of Threats and a Public Dataset. *IEEE Communication Surveys & Tutorials* 18, 1 (2016), 1–163. arxiv:arXiv:1011.1669v3 https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/PSP/asis-colombia-2016.pdf.

[21] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, and P. Faruki. 2019. Network Intrusion Detection for IoT Security Based on Learning Techniques. *IEEE Communications Surveys Tutorials* 21, 3 (thirdquarter 2019), 2671–2701.

[22] Anirban Chakraborty, Manaar Alam, Vishal Dey, Anupam Chattopadhyay, and Debdeep Mukhopadhyay. 2018. Adversarial Attacks and Defences: A Survey. (2018). arxiv:cs.LG/1810.00069 http://arxiv.org/abs/1810.00069

[23] Guillaume Chapron. 2017. The environment needs cryptogovernance. *Nature* 545, 7655 (2017).

[24] Baibhab Chatterjee, Debayan Das, and Shreyas Sen. 2018. RF-PUF: IoT security enhancement through authentication of wireless nodes using in-situ machine learning. *Proceedings of the 2018 IEEE International Symposium on Hardware Oriented Security and Trust, HOST 2018* PP, c (2018), 205–208. arxiv:1805.01048

[25] Konstantinos Christidis and Michael Devetsikiotis. 2016. Blockchains and Smart Contracts for the Internet of Things. *IEEE Access* 4 (2016). http://ieeexplore.ieee.org/document/7467408/

[26] Kelton A. P. da Costa, João P. Papa, Celso O. Lisboa, Roberto Munoz, and Victor Hugo C. de Albuquerque. 2019. Internet of Things: A survey on machine learning-based intrusion detection approaches. *Computer Networks* 151 (2019), 147–157. https://@doi.org/10.1016/j.comnet.2019.01.023

[27] Tim Dalgleish, J. Mark G. Williams, Ann-Marie J. Golden, Nicola Perkins, Lisa Feldman Barrett, Phillip J. Barnard, Cecilia Au Yeung, Victoria Murphy, Rachael Elward, Kate Tchanturia, and Edward Watkins. 2018. The Blockchain-enabled Intelligent IoT Economy. Retrieved from https://www.forbes.com/sites/cognitiveworld/2018/10/04/the-blockchain-enabled-intelligent-iot-economy/#14b65de82a59.

[28] Guido Dartmann, Houbing Song, and Anke Schmeink. 2019. *Big Data Analytics for Cyber-Physical Systems: Machine Learning for the Internet of Things*. Elsevier. 1–360 pages.

[29] Tien Tuan Anh Dinh, Ji Wang, Gang Chen, Rui Liu, Beng Chin Ooi, and Kian-Lee Tan. 2017. BLOCKBENCH: A Framework for Analyzing Private Blockchains. arxiv:cs.DB/1703.04057

[30] Abebe Diro and Naveen Chilamkurti. 2018. Leveraging LSTM Networks for Attack Detection in Fog-to-Things Communications. *IEEE Communications Magazine* 56, 9 (2018).

[31] Ali Dorri, Salil S. Kanhere, and Raja Jurdak. 2016. Blockchain in Internet of Things: Challenges and Solutions. *CoRR* abs/1608.05187 (2016). arxiv:1608.05187 http://arxiv.org/abs/1608.05187.

[32] Ali Dorri, Salil S Kanhere, Raja Jurdak, and Praveen Gauravaram. 2017. Blockchain for IoT Security and Privacy: The Case Study of a Smart Home. In *Proceedings of the 2nd IEEE PERCOM Workshop On Security Privacy And Trust In The Internet of Things 2017 Blockchain*. Hawaii.

[33] Paul Dunphy and Fabien A.P. Petitcolas. 2018. A first look at identity management schemes on the blockchain. *IEEE Security and Privacy* 16, 4 (2018). arxiv:1801.03294

[34] Omar E. Elejla, Bahari Belaton, Mohammed Anbar, Basim Alabsi, and Ahmed K. Al-Ani. 2019. Comparison of classification algorithms on ICMPv6-based DDoS attacks detection. *Lecture Notes in Electrical Engineering* 481 (2019), 347–357.

[35] EMarketer. 2016. Number of Smartphone Users Worldwide From 2014 to 2020 (In Billions). Retrieved from https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/.

[36] Amin Fadaeddini, Babak Majidi, and Mohammad Eshghi. 2020. Secure decentralized peer-to-peer training of deep neural networks based on distributed ledger technology. *The Journal of Supercomputing*0123456789 (2020). https://doi.org/10.1007/s11227-020-03251-9

[37] Kai Fan, Yanhui Ren, Yue Wang, Hui Li, and Yingtang Yang. 2018. Blockchain-based efficient privacy preserving and data sharing scheme of content-centric network in 5G. *IET Communications* 12, 5 (2018). http://digital-library.theiet.org/content/journals/10.1049/iet-com.2017.0619

[38] Kai Fan, Shangyang Wang, Yanhui Ren, Kan Yang, Zheng Yan, Hui Li, and Yintang Yang. 2019. Blockchain-based Secure Time Protection Scheme in IoT. *IEEE Internet of Things Journal* 6, 3 (2019), 4671–4679.

[39] Bo Feng, Qiang Fu, Mianxiong Dong, Dong Guo, and Qiang Li. 2018. Multistage and Elastic Spam Detection in Mobile Social Networks through Deep Learning. *IEEE Network* 32, 4 (2018), 15–21.

[40] P. Feng, J. Ma, C. Sun, X. Xu, and Y. Ma. 2018. A Novel Dynamic Android Malware Detection System With Ensemble Learning. *IEEE Access* 6 (2018), 30996–31011.

[41] M. D. Sadek Ferdous, Mohammad Jabed Morshed Chowdhury, Kamanashis Biswas, Niaz Chowdhury, and Vallipuram Muthukkumarasamy. 2020. Immutable autobiography of smart cars leveraging blockchain technology. *The Knowledge Engineering Review* 22 (2020), e3.

[42] T. M. Fernández-Caramés and P. Fraga-Lamas. 2018. A Review on the Use of Blockchain for the Internet of Things. *IEEE Access* 6 (2018).

[43] Luca Ferretti, Francesco Longo, Michele Colajanni, Giovanni Merlino, and Nachiket Tapas. 2019. Authorization transparency for accountable access to IoT services. In *Proceedings of the 2019 IEEE International Congress on Internet of Things, ICIOT 2019—Part of the 2019 IEEE World Congress on Services* (2019), 91–99.

[44] Jianbin Gao, Kwame Omono Asamoah, Emmanuel Boateng Sifah, Abla Smahi, Qi Xia, Hu Xia, Xiaosong Zhang, and Guishan Dong. 2018. GridMonitoring: Secured Sovereign Blockchain Based Monitoring on Smart Grid. *IEEE Access* 6 (2018).

[45] M. Giles. 2019. Five Emerging Cyber-threats to Worry About in 2019. Retreived from https://www.technologyreview.com/s/612713/five-emerging-cyber-threats-2019/.
[46] Akhil Goel, Akshay Agarwal, Mayank Vatsa, Richa Singh, and Nalini Ratha. 2019. DeepRing: Protecting Deep Neural Network with Blockchain. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2019), 1–8.
[47] A. Goel, A. Singh, A. Agarwal, M. Vatsa, and R. Singh. 2018. SmartBox: Benchmarking Adversarial Detection and Mitigation Algorithms for Face Recognition. In *Proceedings of the 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. 1–7.
[48] Vindu Goel and Nicole Perlroth. 2016. Yahoo Says 1 Billion User Accounts Were Hacked. Retrieved from https://www.nytimes.com/2016/12/14/technology/yahoo-hack.html.
[49] Tomer Golomb, Yisroel Mirsky, and Yuval Elovici. 2018. CIoTA: Collaborative IoT Anomaly Detection via Blockchain. *CoRR* abs/1803.03807 (2018). arxiv:1803.03807 http://arxiv.org/abs/1803.03807.
[50] Gaurav Goswami, Akshay Agarwal, Nalini Ratha, Richa Singh, and Mayank Vatsa. 2019. Detecting and Mitigating Adversarial Perturbations for Robust Face Recognition. *International Journal of Computer Vision* 127, 6–7 (June 2019), 719–742. https://@doi.org/10.1007/s11263-019-01160-w
[51] Gaurav Goswami, Nalini Ratha, Akshay Agarwal, Richa Singh, and Mayank Vatsa. 2018. Unravelling Robustness of Deep Learning based Face Recognition Against Adversarial Attacks. arxiv:cs.CV/1803.00401
[52] Jingjing Gu, Binglin Sun, Xiaojiang Du, and Senior Member. 2018. Consortium Blockchain-Based Malware Detection in Mobile Devices. *IEEE Access* 6 (2018).
[53] Rui Guo, Huixian Shi, Qinglan Zhao, and Dong Zheng. 2018. Secure Attribute-Based Signature Scheme with Multiple Authorities for Blockchain in Electronic Health Records Systems. *IEEE Access* 6 (2018). arxiv:0912.0833
[54] Yash Gupta, Rajeev Shorey, Devadatta Kulkarni, and Jeffrey Tew. 2018. The applicability of blockchain in the Internet of Things. In *Proceedings of the 2018 10th International Conference on Communication Systems and Networks, (COMSNETS 2018)* .
[55] Haya R. Hasan and Khaled Salah. 2018. Blockchain-Based Proof of Delivery of Physical Assets with Single and Multiple Transporters. *IEEE Access* 6 (2018).
[56] Muneeb Ul Hassan, Mubashir Husain Rehmani, and Jinjun Chen. 2019. Privacy preservation in blockchain based IoT systems: Integration issues, prospects, challenges, and future research directions. *Future Generation Computer Systems* 97 (2019), 512–529. https://@doi.org/10.1016/j.future.2019.02.060
[57] Vikas Hassija, Vinay Chamola, Vikas Saxena, Divyansh Jain, Pranav Goyal, and Biplab Sikdar. 2019. A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures. *IEEE Access* 7 (2019), 82721–82743.
[58] Fatima Hussain, Syed Ali Hassan, Rasheed Hussain, and Ekram Hossain. 2020. Machine Learning for Resource Management in Cellular and IoT Networks: Potentials, Current Solutions, and Open Challenges. *IEEE Communications Surveys & Tutorials*c (2020), 1–1. arxiv:1907.08965
[59] F. Hussain, R. Hussain, S. A. Hassan, and E. Hossain. 2020. Machine Learning in IoT Security: Current Solutions and Future Challenges. *IEEE Communications Surveys Tutorials* (2020), 1–1.
[60] Alfonso Panarello Id and Nachiket Tapas. 2018. *Blockchain and IoT Integration : A Systematic Survey*.
[61] Muhammad Ikram, Pierrick Beaume, and Mohamed Ali Kâafar. 2019. DaDiDroid: An Obfuscation Resilient Tool for Detecting Android Malware via Weighted Directed Call Graph Modelling. In *Proceedings of the 16th International Joint Conference on e-Business and Telecommunications, ICETE 2019 - Volume 2: SECRYPT, Prague, Czech Republic, July 26-28, 2019*. SciTePress, 211–219. https://@doi.org/10.5220/0007834602110219
[62] Muhammad Ikram, Rahat Masood, Gareth Tyson, Mohamed Ali Kaafar, Noha Loizon, and Roya Ensafi. 2019. The chain of implicit trust: An analysis of the web third-party resources loading. In *The World Wide Web Conference*.
[63] Ahmad Javaid, Quamar Niyaz, Weiqing Sun, and Mansoor Alam. 2016. A Deep Learning Approach for Network Intrusion Detection System. *EAI Endorsed Transactions on Security and Safety* 3, 9 (5 2016).
[64] Sabina Jeschke, Christian Brecher, Houbing Song, and Danda B. Rawat. 2017. *Industrial Internet of Things: Cybermanufacturing Systems*. Springer. 1–715 pages.
[65] Zhanglong Ji, Zachary Chase Lipton, and Charles Elkan. 2014. Differential Privacy and Machine Learning: a Survey and Review. *CoRR* abs/1412.7584 (2014). arxiv:1412.7584 http://arxiv.org/abs/1412.7584
[66] Qi Jia, Linke Guo, Zhanpeng Jin, and Yuguang Fang. 2018. Preserving model privacy for machine learning in distributed systems. *IEEE Transactions on Parallel and Distributed Systems* 29, 8 (2018), 1808–1822.
[67] Qi Jing, Athanasios Vasilakos, Jiafu Wan, Jingwei Lu, and Dechao Qiu. 2014. Security of the Internet of Things: Perspectives and challenges. *Wireless Networks* 20 (11 2014), 2481–2501.
[68] Xuyang Jing, Zheng Yan, Xueqin Jiang, and Witold Pedrycz. 2019. Network traffic fusion and analysis against DDoS flooding attacks with a novel reversible sketch. *Information Fusion* 51 (2019), 100–113. https://@doi.org/10.1016/j.inffus.2018.10.013

[69] Jiawen Kang, Rong Yu, Xumin Huang, Sabita Maharjan, Yan Zhang, and Ekram Hossain. 2017. Enabling Localized Peer-to-Peer Electricity Trading among Plug-in Hybrid Electric Vehicles Using Consortium Blockchains. *IEEE Transactions on Industrial Informatics* 13, 6 (2017).

[70] Jiawen Kang, Rong Yu, Xumin Huang, Maoqiang Wu, Sabita Maharjan, Shengli Xie, and Yan Zhang. 2018. Blockchain for Secure and Efficient Data Sharing in Vehicular Edge Computing and Networks. *IEEE Internet of Things Journal* (2018). https://ieeexplore.ieee.org/document/8489897/

[71] M. Kang and J. Kang. 2016. A Novel Intrusion Detection Method Using Deep Neural Network for In-Vehicle Network Security. In *Proceedings of the 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. 1–5.

[72] P. Kasinathan, C. Pastrone, M. A. Spirito, and M. Vinkovits. 2013. Denial-of-Service detection in 6LoWPAN based Internet of Things. In *Proceedings of the 2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. 600–607.

[73] Minhaj Ahmad Khan and Khaled Salah. 2018. IoT security: Review, blockchain solutions, and open challenges. *Future Generation Computer Systems* 82 (2018). arxiv:1705.08230 https://@doi.org/10.1016/j.future.2017.11.022

[74] Djamel Eddine Kouicem, Abdelmadjid Bouabdallah, and Hicham Lakhlef. 2018. Internet of Things security: A top-down survey. *Computer Networks* 141 (2018), 199–221. https://@doi.org/10.1016/j.comnet.2018.03.012

[75] Nir Kshetri. 2017. Blockchain's roles in strengthening cybersecurity and protecting privacy. *Telecommunications Policy* 41, 10 (2017). https://@doi.org/10.1016/j.telpol.2017.09.003

[76] Nallapaneni Manoj Kumar and Pradeep Kumar Mallick. 2018. Blockchain technology for security issues and challenges in IoT. *Procedia Computer Science* 132 (2018), 1815–1823. https://@doi.org/10.1016/j.procs.2018.05.140

[77] Boohyung Lee and Jong Hyouk Lee. 2017. Blockchain-based secure firmware update for embedded devices in an Internet of Things environment. *Journal of Supercomputing* 73, 3 (2017).

[78] Jong Hyouk Lee. 2017. BIDaaS: Blockchain based ID as a service. *IEEE Access* 6 (2017).

[79] Lun Li, Jiqiang Liu, Lichen Cheng, Shuo Qiu, Wei Wang, Xiangliang Zhang, and Zonghua Zhang. 2018. CreditCoin: A privacy-preserving blockchain-based incentive announcement network for communications of smart vehicles. *IEEE Transactions on Intelligent Transportation Systems* 19, 7 (2018).

[80] Xiaoqi Li, Peng Jiang, Ting Chen, Xiapu Luo, and Qiaoyan Wen. 2017. A survey on the security of blockchain systems. *Future Generation Computer Systems* (2017). arxiv:1802.06993 http://dx.@doi.org/10.1016/j.future.2017.08.020

[81] Yuancheng Li, Rong Ma, and Runhai Jiao. 2015. A hybrid malicious code detection method based on deep learning. *International Journal of Software Engineering and Its Applications* 9, 205–216.

[82] Fan Liang, William Grant Hatcher, Weixian Liao, Weichao Gao, and Wei Yu. 2019. Machine learning for security and the Internet of Things: The good, the bad, and the ugly. *IEEE Access* 7 (2019), 158126–158147.

[83] X. Liang, J. Zhao, S. Shetty, and D. Li. 2017. Towards data assurance and resilience in IoT using blockchain. In *2017 IEEE Military Communications Conference (MILCOM 2017)*. 261–266.

[84] Qiang Liu, Pan Li, Wentao Zhao, and Wei Cai. 2018. A survey on security threats and defensive techniques of machine learning : A data driven view. *IEEE Access* 6 (2018), 12103–12117.

[85] Zhaojun Lu, Wenchao Liu, Qian Wang, Gang Qu, and Zhenglin Liu. 2018. A privacy-preserving trust model based on blockchain for VANETs. *IEEE Access* 6 (2018).

[86] Xindi Ma, Jianfeng Ma, Hui Li, Qi Jiang, and Sheng Gao. 2018. PDLM: Privacy-preserving deep learning model on cloud with multiple keys. *IEEE Transactions on Services Computing* (2018), 1–13.

[87] Caciano Machado and Antonio Augusto Frohlich. 2018. IoT data integrity verification for cyber-physical systems using blockchain. *Proceedings - 2018 IEEE 21st International Symposium on Real-Time Computing, ISORC 2018* (2018).

[88] Lorenzo Fernández Maimó, Ángel Luis, Perales Gómez, Félix J. García Clemente, Manuel G. I. L. Pérez, and Gregorio Martínez Pérez. 2018. A self-adaptive deep learning-based system for anomaly detection in 5G networks. *IEEE Access* 6 (2018).

[89] I. Makhdoom, M. Abolhasan, J. Lipman, R. P. Liu, and W. Ni. 2019. Anatomy of threats to the Internet of Things. *IEEE Communications Surveys Tutorials* 21, 2 (Secondquarter 2019).

[90] Polina Mamoshina, Lucy Ojomoko, Yury Yanovich, Alex Ostrovski, Alex Botezatu, Pavel Prikhodko, Evgeny Izumchenko, Alexander Aliper, Konstantin Romantsov, Alexander Zhebrak, Iraneus Ogu, and Alexander Zhavoronkov. 2018. Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare. *Oncotarget* 9 (01 2018), 5665–5690.

[91] Rahat Masood, Dinusha Vatsalan, Muhammad Ikram, and Mohamed Ali Kaafar. 2018. Incognito: A method for obfuscating web data. In *Proceedings of the 2018 World Wide Web Conference*.

[92] Gihan J. Mendis, Moein Sabounchi, Jin Wei, and Rigoberto Roche'. 2018. Blockchain as a service: An autonomous, privacy preserving, decentralized architecture for deep learning. *CoRR* abs/1807.02515 (2018). arxiv:1807.02515 http://arxiv.org/abs/1807.02515

[93] Gihan J. Mendis, Yifu Wu, Jin Wei, Moein Sabounchi, and Rigoberto Roche. 2020. A blockchain-powered decentralized and secure computing paradigm. *IEEE Transactions on Emerging Topics in Computing* (2020), 1–18. http://dx.doi.org/10.1109/TETC.2020.2983007

[94] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han. 2018. When intrusion detection meets blockchain technology: A review. *IEEE Access* 6 (2018), 10179–10188.

[95] Jelena Milosevic, Miroslaw Malek, and Alberto Ferrante. 2016. A friend or a Foe? Detecting malware using memory and CPU features. In *Proceedings of the 13th International Joint Conference on e-Business and Telecommunications (ICETE 2016)*, Vol. 4. 73–84.

[96] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli. 2019. A detailed investigation and analysis of using machine learning techniques for intrusion detection. *IEEE Communications Surveys Tutorials* 21, 1 (Firstquarter 2019), 686–728.

[97] N. Moustafa, B. Turnbull, and K. R. Choo. 2019. An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of Internet of Things. *IEEE Internet of Things Journal* 6, 3 (June 2019), 4815–4830.

[98] H. Niwa. 2007. Why Blockchain is the Future of IoT? Retrieved from https://www.networkworld.com/article/3200029/internet-of-things/why-blockchain-is-the-future-of-iot.html.

[99] Andrea Peterson. 2014. eBay Asks 145 Million Users to Change Passwords After Data Breach. Retrieved from https://www.washingtonpost.com/news/the-switch/wp/2014/05/21/ebay-asks-145-million-users-to-change-passwords-after-data-breach/.

[100] Cookies Policy. 2019. WhatsApp hack: Is any app or computer truly secure? - BBC News. Accessed on September 16, 2020 from https://www.bbc.com/news/technology-48282092.

[101] S. Prabavathy, K. Sundarakantham, and S. Mercy Shalinie. 2018. Design of cognitive fog computing for intrusion detection in Internet of Things. *Journal of Communications and Networks* 20, 3 (2018), 291–298.

[102] Pavithra Prabhu and K. N. Manjunath. 2019. Secured image transmission in medical imaging applications—a survey. In *Computer Aided Intervention and Diagnostics in Clinical and Medical Images*. Springer International Publishing, Cham, 125–133.

[103] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. 1992. *Numerical Recipes in C (2nd Ed.): The Art of Scientific Computing*. Cambridge University Press, New York, NY.

[104] W. Nicholson Price and I. Glenn Cohen. 2019. Privacy in the age of medical big data. *Nature Medicine* 25, 1 (2019), 37–43. arxiv:arXiv:1011.1669v3 http://dx.@doi.org/10.1038/s41591-018-0272-7

[105] Yogachandran Rahulamathavan, Raphael C. Phan, Sudip Misra, and Muttukrishnan Rajarajan. 2017. Privacy-preserving Blockchain based IoT Ecosystem using Attribute-based Encryption. In *Proceedings of the 2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, October (2017).

[106] Francesco Restuccia, Salvatore DrOro, and Tommaso Melodia. 2018. Securing the Internet of Things in the age of machine learning and software-defined networking. *IEEE Internet of Things Journal* 1, 1 (2018). arxiv:1803.05022

[107] Ana Reyna, Cristian Martín, Jaime Chen, Enrique Soler, and Manuel Díaz. 2018. On blockchain and its integration with IoT. Challenges and opportunities. *Future Generation Computer Systems* 88 (2018). arxiv:arXiv:1011.1669v3

[108] Mostafa Rezazad, Matthias R. Brust, Mohammad Akbari, Pascal Bouvry, and Ngai-Man Cheung. 2018. Detecting target-area link-flooding DDoS attacks using traffic analysis and supervised learning. *Advances in Information and Communication Networks*, 180–202. http://dx.@doi.org/10.1007/978-3-030-03405-4_12

[109] Mayu Sakurada and Takehisa Yairi. 2014. Anomaly detection using autoencoders with nonlinear dimensionality reduction. In *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis (MLSDA'14)*. ACM, New York, NY, Article 4, 8 pages. http://@doi.acm.org/10.1145/2689746.2689747

[110] Khaled Salah, M. Habib Ur Rehman, Nishara Nizamuddin, and Ala Al-Fuqaha. 2019. Blockchain for AI: Review and open research challenges. *IEEE Access* 7 (2019), 10127–10149.

[111] T. Salman, M. Zolanvari, A. Erbad, R. Jain, and M. Samaka. 2019. Security services using blockchains: A state of the art survey. *IEEE Communications Surveys Tutorials* 21, 1 (Firstquarter 2019), 858–880.

[112] Pradip Kumar Sharma, Mu Yen Chen, and Jong Hyuk Park. 2018. A software defined fog node based distributed blockchain cloud architecture for IoT. *IEEE Access* 6 (2018), 115–124.

[113] Pradip Kumar Sharma, Saurabh Singh, Young Sik Jeong, and Jong Hyuk Park. 2017. DistBlockNet: A distributed blockchains-based secure SDN architecture for IoT networks. *IEEE Communications Magazine* 55, 9 (2017), 78–85.

[114] V. Sharma, I. You, F. Palmieri, D. N. K. Jayakody, and J. Li. 2018. Secure and energy-effcient handover in fog networks using blockchain-based DMM. *IEEE Communications Magazine* 56, 5 (2018), 22–31.

[115] Shaila Sharmeen, Shamsul Huda, Jemal H. Abawajy, Walaa Nagy Ismail, and Mohammad Mehedi Hassan. 2018. Malware threats and detection for industrial mobile-IoT networks. *IEEE Access* 6 (2018), 15941–15957.

[116] M. Shen, X. Tang, L. Zhu, X. Du, and M. Guizani. 2019. Privacy-preserving support vector machine training over blockchain-based encrypted IoT data in smart cities. *IEEE Internet of Things Journal* 6, 5 (2019), 7702–7712.

[117] IEEE Computer Society, Institute of Electrical, and Electronics Engineers. 2005. *IEEE Annals of the History of Computing*. Number v. 27-28. IEEE Computer Society. 92650021 https://books.google.com.au/books?id=xv9UAAAAMAAJ

[118] Houbing Song, Glenn Fink, and Sabina Jeschke. 2017. *Security and Privacy in Cyber-Physical Systems: Foundations, Principles, and Applications*. Wiley Online Library. 1–472 pages.

[119] Juah C. Song, Mevlut A. Demir, John J. Prevost, and Paul Rad. 2018. Blockchain design for trusted decentralized IoT networks. In *Proceedings of the 2018 13th System of Systems Engineering Conference (SoSE 2018)*.

[120] Tianyi Song, Ruinian Li, Bo Mei, Jiguo Yu, Xiaoshuang Xing, and Xiuzhen Cheng. 2018. A privacy preserving communication protocol for IoT applications in smart homes. In *Proceedings of the 2016 International Conference on Identification, Information and Knowledge in the Internet of Things (IIKI 2016)* 519–524.

[121] Xiaoqiang Sun, Peng Zhang, Joseph K. Liu, Jianping Yu, and Weixin Xie. 2018. Private machine learning classification based on fully homomorphic encryption. *IEEE Transactions on Emerging Topics in Computing* 6750, c (2018).

[122] Dan Swinhoe. 2019. What is a Man-in-the-middle Attack? How MitM Attacks Work and How to Prevent Them. Retrieved from https://www.csoonline.com/article/3340117/what-is-a-man-in-the-middle-attack-how-mitm-attacks-work-and-how-to-prevent-them.html.

[123] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R. P. Liu. 2014. A system for denial-of-service attack detection based on multivariate correlation analysis. *IEEE Transactions on Parallel and Distributed Systems* 25, 2 (Feb. 2014), 447–456.

[124] Z. Tan, A. Jamdagni, X. He, P. Nanda, R. P. Liu, and J. Hu. 2015. Detection of denial-of-service attacks based on computer vision techniques. *IEEE Transactions on Computing* 64, 9 (Sep. 2015), 2519–2533.

[125] Zhushou Tang, Ke Tang, Minhui Xue, Yuan Tian, Sen Chen, Muhammad Ikram, Tielei Wang, and Haojin Zhu. 2020. iOS, your OS, everybody's OS: Vetting and analyzing network services of iOS applications. In *Proceedings of the 29th USENIX Security Symposium (USENIX Security 2020)*. USENIX Association, Boston, MA.

[126] N. Tapas, G. Merlino, and F. Longo. 2018. Blockchain-based IoT-cloud authorization and delegation. In *Proceedings of the 2018 IEEE International Conference on Smart Computing (SMARTCOMP)*. 411–416.

[127] Florian Tschorsch and Björn Scheuermann. 2016. Bitcoin and beyond: A technical survey on decentralized digital currencies. *IEEE Communication Surveys & Tutorials* 18, 3 (2016), 2084–2123.

[128] C. Tselios, I. Politis, and S. Kotsopoulos. 2017. Enhancing SDN security for iot-related deployments through blockchain. In *Proceedings of the 2017 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN 2017)* .

[129] A. ul Haque, M. S. Ghani, and T. Mahmood. 2020. Decentralized transfer learning using blockchain IPFS for deep learning. In *Proceedings of the 2020 International Conference on Information Networking (ICOIN)*. 170–177.

[130] Jingjun Wang, Shengshan Hu, Qian Wang, and Yutao Ma. 2017. Privacy-preserving outsourced feature extractions in the cloud: A survey. *IEEE Network* October (2017), 36–41.

[131] Jingzhong Wang, Mengru Li, Yunhua He, Hong Li, Ke Xiao, and Chao Wang. 2018. A blockchain based privacy-preserving incentive mechanism in crowdsensing applications. *IEEE Access* 6 (2018), 17545–17556.

[132] Ning Wang, Ting Jiang, Shichao Lv, and Liang Xiao. 2017. Physical-layer authentication based on extreme learning machine. *IEEE Communications Letters* 21, 7 (2017), 1557–1560.

[133] Wei Wang, Zhenzhen Gao, Meichen Zhao, Yidong Li, Jiqiang Liu, and Xiangliang Zhang. 2018. DroidEnsemble: Detecting android malicious applications with ensemble of string and structural static features. *IEEE Access* 6 (2018), 31798–31807.

[134] Xu Wang, Xuan Zha, Wei Ni, Ren Ping Liu, Y. Jay Guo, Xinxin Niu, and Kangfeng Zheng. 2019. Survey on blockchain for Internet of Things. *Computer Communications* 136, January (2019), 10–29. https://@doi.org/10.1016/j.comcom.2019.01.006

[135] Linfeng Wei, Weiqi Luo, Jian Weng, Yanjun Zhong, Xiaoqian Zhang, and Zheng Yan. 2017. Machine learning-based malicious application detection of android. *IEEE Access* 5 (2017), 25591–25601.

[136] Jiasi Weng, Jian Weng, Jilian Zhang, Ming Li, Yue Zhang, and Weiqi Luo. 2019. DeepChain: Auditable and privacy-preserving deep learning with blockchain-based incentive. *IEEE Transactions on Dependable and Secure Computing* (2019), 1–18.

[137] Liang Xiao, Yan Li, and Guoan Han. 2016. PHY-layer spoofing detection with reinforcement learning in wireless networks. *IEEE Transactions on Vehicular Technology* 65, 12 (2016), 10037–10047.

[138] Liang Xiao, Donghua Jiang, Dongjin Xu, and Ning An. 2018. Secure mobile crowdsensing with deep learning. *China Communications* 15 (2018). arxiv:1801.07379 http://arxiv.org/abs/1801.07379

[139] Liang Xiao, Xiaoyue Wan, Xiaozhen Lu, Yanyong Zhang, and Di Wu. 2018. IoT security techniques based on machine learning: How do IoT devices use AI to enhance security? *IEEE Signal Processing Magazine* 35, 5 (2018). arxiv:1801.06275

[140] Weizhong Yan and Lijie Yu. 2015. On accurate and reliable anomaly detection for gas turbine combustors: A deep learning approach.

[141] Bin Yu, Jarod Wright, Surya Nepal, Liming Zhu, Joseph Liu, and Rajiv Ranjan. 2018. IoTChain: Establishing trust in the Internet of Things ecosystem using blockchain. *IEEE Cloud Computing* 5, 4 (2018). https://ieeexplore.ieee.org/document/8436081/

[142] Kevin Kam Fung Yuen. 2019. Towards a cybersecurity investment assessment method using primitive cognitive network process. In *Proceedings of the 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. 068–071.

[143] Tao Zhang and Quanyan Zhu. 2018. Distributed privacy-preserving collaborative intrusion detection systems for VANETs. *IEEE Transactions on Signal and Information Processing Over Networks* 4, 1 (2018), 148–161.

[144] Benjamin Zi Hao Zhao, Muhammad Ikram, Hassan Jameel Asghar, Mohamed Ali Kaafar, Abdelberi Chaabane, and Kanchana Thilakarathna. 2019. A decade of mal-activity reporting: A retrospective analysis of internet malicious activity blacklists. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*. 193–205.

[145] Lijing Zhou, Licheng Wang, Yiru Sun, and Pin Lv. 2018. BeeKeeper: A blockchain-based IoT system with secure storage and homomorphic computation. *IEEE Access* 6 (2018).

[146] Yiyun Zhou, Meng Han, Liyuan Liu, Jing Selena He, and Yan Wang. 2018. Deep learning approach for cyberattack detection. In *Proceedings of the IEEE Conference on Computer Communications Workshops* (INFOCOM 2018).

[147] Hui Zhu, Xiaoxia Liu, Rongxing Lu, and Hui Li. 2017. Efficient and privacy-preserving online medical prediagnosis framework using nonlinear SVM. *IEEE Journal of Biomedical and Health Informatics* 21, 3 (2017), 838–850.

[148] Xiaoyang Zhu and Youakim Badr. 2018. Identity management systems for the Internet of Things: A survey towards blockchain solutions. *Sensors (Basel, Switzerland)* 18, 12 (2018), 1–18.

Received December 2019; revised June 2020; accepted July 2020