

Received August 2, 2020, accepted August 14, 2020, date of publication August 20, 2020, date of current version September 1, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3018170

Research on Artificial Intelligence Enhancing Internet of Things Security: A Survey

HUI WU¹, HAITING HAN^{ID2}, (Member, IEEE), XIAO WANG¹, AND SHENGLI SUN¹

¹School of Software and Microelectronics, Peking University, Beijing 100181, China

²Department of Food and Resource Economics, University of Copenhagen, 1958 Copenhagen, Denmark

Corresponding author: Shengli Sun (slsun@ss.pku.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFB1402900 and Grant 2018YFB1403000, and in part by the Natural Science Foundation of Jiangsu Province under Grant BK20151132.

ABSTRACT Through three development routes of authentication, communication, and computing, the Internet of Things (IoT) has become a variety of innovative integrated solutions for specific applications. However, due to the openness, extensiveness and resource constraints of IoT, each layer of the three-tier IoT architecture suffers from a variety of security threats. In this work, we systematically review the particularity and complexity of IoT security protection, and then find that Artificial Intelligence (AI) methods such as Machine Learning (ML) and Deep Learning (DL) can provide new powerful capabilities to meet the security requirements of IoT. We analyze the technical feasibility of AI in solving IoT security problems and summarize a general process of AI solutions for IoT security. For four serious IoT security threats: device authentication, Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks defense, intrusion detection and malware detection, we summarize representative AI solutions and compare the different algorithms and technologies used by various solutions. It should be noted that although AI provides many new capabilities for the security protection of IoT, it also brings new potential challenges and possible negative effects to IoT in terms of data, algorithm and architecture. In the future, how to solve these challenges can serve as potential research directions.

INDEX TERMS Artificial intelligence, deep learning, Internet of Things, machine learning, security.

I. INTRODUCTION

The International Telecommunication Union formally proposed the concept of “Internet of Things” at the World Summit on the Information Society (WSIS) in 2005 [1]. The Internet of things (IoT) refers to a distributed network that combines various sensor devices and systems, such as sensor networks, RFID devices, barcode and QR code devices, global positioning systems, etc. [2], with the Internet through wired and wireless communication technologies, enabling embedded systems to communicate and interconnect.

From the discovery of electromagnetic induction to RFID, from simple sensors to ubiquitous connections, from electronic toll collection (ETC) to smart cities, the development of IoT has always been along the following three technical routes:

- *The development of sensing, identification and authentication technologies.* Sensing, identification and

authentication technologies are the foundation of IoT. As nerve endings of IoT, sensors are the largest and most basic part of the chain of IoT. A large number of general-purpose sensor devices have been popularized, and high-end sensor devices in specific fields have also made great progress.

- *The development of transmission and communication technologies.* Transmission and communication technologies are the guarantee of IoT. The large amount of information collected by IoT devices needs to be transmitted and aggregated to the central node or the processing unit in a more convenient, more reliable, and safer way. The development of wired and wireless networks, cellular networks, and other transmission and communication technologies have made it possible for large-scale IoT data transmission.
- *The development of data computing and processing technologies.* Data computing and processing technologies are essential to provide applications and services using IoT data. IoT applications need real-time

The associate editor coordinating the review of this manuscript and approving it for publication was Rongxing Lu^{ID}.

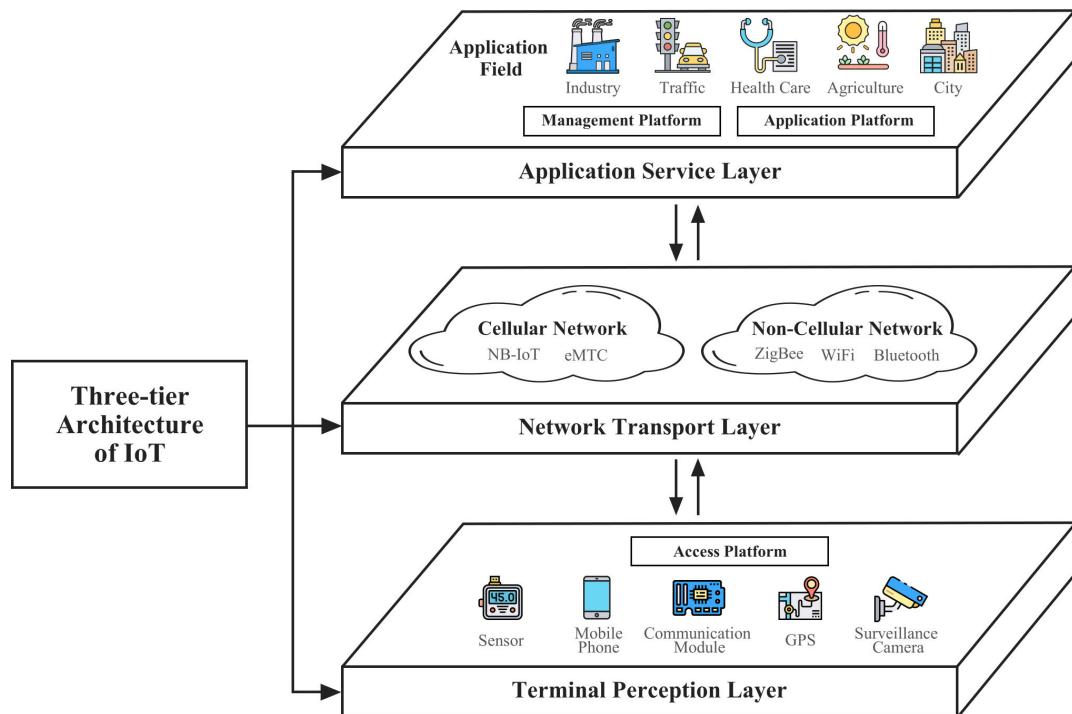


FIGURE 1. Three-tier architecture of IoT.

perception and intelligent feedback for a large number of information nodes. The development of data computing and processing technologies is the key to improve processing intelligence and effectiveness. Computing technologies such as artificial intelligence and cloud computing have given IoT more possibilities to provide advanced application services.

With the development of technologies and the expansion of application fields, IoT has gradually evolved into a set of solutions for specific applications. IoT focuses on the integration and innovation of solutions and will become a new idea, new tool and new method of social governance, combining the Internet with the physical world and providing intelligent interaction. The numerous applications and services provided by IoT cover many fields such as manufacturing, energy management (such as smart grids), urban life (such as smart cities), and personal healthcare.

According to the entire process of information generation, transmission and processing, referring to the traditional architecture [3] and ISO/IEC 30141:2018 “IoT Reference Architecture” [4], IoT generally has an entity-based architecture that can divide IoT from bottom to top into three layers, namely terminal perception layer, network transport layer, and application service layer. The specific architecture is shown in Fig. 1. This architecture integrates various entities involved in the IoT and shows that they have an interactive relationship with each other.

The terminal perception layer is the source of IoT data collection. The collected information of various objects will be transmitted to the upper layer. The entities involved in

terminal perception layer include physical entities representing real things, IoT device entities (sensor devices, identification devices represented by RFID, and positioning / tracking devices represented by GPS), and access platform connecting local IoT device network and wide area network.

The network transport layer transmits information from the terminal perception layer to the application service layer to realize the communication and connection functions. The network transport layer uses non-cellular networks (ZigBee, Bluetooth and Wi-Fi, etc.) and cellular networks (NB-IoT, eMTC, etc.) for data encoding, authentication, and transmission.

The application service layer processes the data transmitted from the network transport layer and integrates them with various industries to support vertical applications of IoT, providing rich and specific services for different users in specific fields, such as smart grids, smart homes, and smart cities. The application service layer also includes application & service subsystems that provide capabilities of data storage, analysis and service, as well as operation maintenance & management subsystems that provide capabilities of management and operation support.

II. IoT SECURITY THREATS

As more and more machines and smart devices are connected to the network, the vulnerabilities of IoT security are gradually exposed. IoT devices are more vulnerable to be attacked than computers or mobile phones, not only because of the surge in the use of IoT devices, but also on account of the complexity, diversity, and inherent mobility of such device

application scenarios. At the same time, IoT has developed rapidly but has not yet matured. The privacy protection crises caused by the openness of the network and the mobility of data are less discussed and regulated. Comprehensive perception makes the data collected and exchanged by IoT more private and dangerous than the Internet.

A. ATTRIBUTION ANALYSIS OF IoT SECURITY THREATS

The widespread popularity and the large-scale deployment have promoted the development of IoT, but also brought new security challenges. Maintaining its security is a complex and challenging task. The reasons for the increasingly serious security problems of IoT are as follows:

The lack of human supervision. IoT terminals are usually deployed in complex and changeable environments to collect information and provide data for applications. However, under these environmental conditions, due to the limitation of human resources, the terminals are exposed, distributed and unattended, so that intruders can easily physically damage devices [5]. Common physical attacks include illegal theft, malicious movement, etc. These attacks will cause damage, data loss and function failure of IoT devices. In the case of huge amount of IoT devices, it is difficult to find and repair damaged terminals in time, which further aggravates the consequences of physical attacks.

The resource constraints of low-power devices and terminals. IoT devices are small in size and low in power consumption. They can do some simple data calculations and are suitable for distributed computing. In recent years, the rapid development of edge computing takes advantage of this characteristic of IoT devices [6]. However, the limited computing capacity and power supply cannot support a large number of complex calculations. There are no remaining resources to implement more fine-grained security measures, resulting in the inability of IoT devices and systems to use complex security mechanisms [7]. The use of some measures may reduce the equipment processing efficiency and increase resource consumption, thus causing damage to the original services. For example, RSA, a commonly used encryption protocol, will consume a lot of resources when running on devices with limited computing power, which is easy to burden the devices. When performing multiple encryption operations at the same time, such as in the Internet of vehicles, the resource consumption will be more serious [8].

The openness of IoT and the expansion of attack scale. Openness is reflected in the various processes of IoT system. IoT can obtain data from various fields, integrate various communication technologies and standards, and provide open services for users in various fields. Openness is conducive to the development of IoT, but also brings the expansion of the scale of potential risks. Due to the interconnection and interdependence of IoT system, any vulnerability can be exploited by attackers to launch large-scale and systematic attacks, which will paralyze the whole system. For example, attackers can use some terminals as the entrance of attack penetration and use tools to analyze the information stored in the

same type of terminal, such as source code and authentication mechanism, so as to invade the whole system.

The integrity and unity of IoT system. IoT combines machines, network infrastructures and application platforms into a complex system. This kind of interconnected system makes operation maintenance and service provision depend on each other and interfere with each other. Although IoT has layered architectures, security problems in all layers are not independent of each other. The problems of one layer may affect other parts of IoT system. If intruders manipulate some terminals to launch DDoS attacks, the whole system may be infected, affecting the application service layer and causing the service to crash; at the same time, attacks against the application platform and software will also cause the leakage of user privacy and malicious manipulation of devices, leading to the abnormality of the terminal perception layer.

The lag of legal supervision and management. IoT security needs not only technologies such as network security, application security or data security, but also needs legal restraints and supervisions of regulatory agencies. However, the rapid application of IoT does not match with the implementation and improvement of its safety supervision mechanism. In recent years, many countries have made laws and standards related to the security of IoT. For example, the US House of Representatives passed the “IoT Cybersecurity Improvement Act 2019”, and the European Telecommunications Standards Institute (ETSI) released the standard for “Consumer IoT Cybersecurity” (TS 103 645) [9]. However, the above-mentioned laws and standards are still in the exploratory stage, and large-scale and industrialized safety management systems have not yet been formed.

B. SECURITY THREATS IN THE TERMINAL PERCEPTION LAYER

Sensors are the main components of the terminal system. Their main function is to monitor objects in real time and collect information. These tiny physical devices are spread over a variety of related engineering fields and its number is extremely large. Most of them are limited by resources, which makes them become the potential attack surface of attackers.

1) THE FIRST PROBLEM FACED BY THE DEVICES IN THE TERMINAL PERCEPTION LAYER IS THE UNPREDICTABLE PHYSICAL ATTACK [5]

Although IoT devices are assets, they are often in a state of lack of supervision. Criminals use theft, damage and other physical means to destroy the connection between devices and central server. For example, in 2018, Chinese sharing economy enterprises (sharing portable chargers and bicycles) suffered large losses or even went bankrupt. A large part of the reason is that smart locks and positioning devices on bicycles and chargers were removed by violence, resulting in asset losses.

2) THERE ARE RISKS OF BEING ATTACKED DURING THE INFORMATION TRANSMISSION OF IoT NODES

There are three main types of terminal perception layer nodes: collection endpoints, information aggregation nodes, and isolated nodes. The collection endpoint mainly corresponds to sensors, which is responsible for sensing and collecting information; the information aggregation node is the server responsible for receiving, processing, forwarding information; the isolation node is embedded equipment responsible for the operations of information encryption and decryption, internal and external network isolation. When information is interconnected between nodes, due to the transmission distance, there are threats such as interception, eavesdropping, counterfeiting, and tampering of nodes.

3) THE IDENTIFICATION AND AUTHENTICATION TECHNOLOGIES ARE INDISPENSABLE PREREQUISITES FOR THE SECURE COMMUNICATION OF IoT DEVICES [10]

Although the uniqueness and certainty of identity can effectively increase the security of IoT devices, hackers can use some ways to bypass this process to implement intrusion. For example, in April 2019, a software called iLnkP2P was discovered without any authentication or encryption measures. Attackers can bypass the firewall with some specific serial numbers and directly establish connections with IoT devices, send malicious messages instead of any valid messages sent by the device.

C. SECURITY THREATS IN THE NETWORK TRANSPORT LAYER

IoT integrates sensor networks [11] and communication networks to form a large-scale network. Similar to the risks faced by the terminal perception layer, the possible attacks also increase significantly with the increase of network scale.

1) THE CONFIDENTIALITY, INTEGRITY, AND AVAILABILITY OF NETWORK ARE TARGETS OF NETWORK TRANSPORT LAYER'S ATTACKERS [12]

Some network targets have poor protection, which makes it easier for attackers to invade. When the system lacks protection and verification mechanisms, the attacker will tamper with the platform software and hardware modules, resulting in the risk of leakage of stored information. Therefore, timely detection of intrusions is critical to curb attacks and protect network security in the early stages.

2) NETWORK TRANSPORT LAYER WILL SUFFER FROM ATTACKS SUCH AS DENIAL OF SERVICE (DoS) AND DISTRIBUTED DENIAL OF SERVICE (DDoS)

Attackers launch DoS and DDoS attacks by sending traffic beyond the target's processing capacity to consume computing and network resources of the target, resulting in resource depletion, thus blocking the target network and causing denial of service. Large-scale DoS and DDoS attacks will cause disastrous consequences to the whole network. Mirai, the botnet

which broke out in 2016, launched a large-scale DDoS attack by using IoT devices, resulting in more than 100,000 devices infected [13].

3) THE COMMUNICATION TECHNOLOGIES USED BY IoT HAVE LIMITATIONS

IoT uses different communication technologies [14], including long-distance networks (NB-IoT and LoRa (Long Range Radio) [15]), short-distance networks (ZigBee [16], Wi-Fi, etc.), and Internet. The security shortcomings of these technologies have been inherited into IoT. For example, the Internet provides a wide range of services for different participants, including IoT users, but at the same time, the communication infrastructure based on TCP / IP is not only vulnerable to security and privacy threats, such as intrusion, replay attacks, and identities theft [17], but also faces challenges such as poor scalability, high complexity, and insufficient resource utilization [18].

D. SECURITY THREATS IN THE APPLICATION SERVICE LAYER

The application service layer processes the data transmitted from the network transport layer and provides services for different application scenarios according to user requirements. Users can directly enjoy the services provided by IoT system through web applications or mobile apps, and enjoy the convenience brought by IoT system. However, there will be system security, data security, and software security problems launched by application-level attackers.

1) SYSTEM SECURITY

The application service layer usually consists of basic environments, components, and virtualized cloud platforms. Basic environments and components, such as operating systems, databases and middleware, will be used by attackers to launch brute force attacks and man-in-the-middle attacks, resulting in unauthorized access, remote control and data leakage. Most IoT systems build virtualized cloud platforms to reduce equipment deployment costs and improve computing performance or business throughput. However, virtualization technology also brings security risks, leading to blurring of the boundary between users and data, resulting in security issues including virtual machine escape, virtual network attacks, and virtualized software vulnerabilities.

2) DATA SECURITY

Databases face security problems. Common database attacks include SQL injection [19], privilege promotion and backup theft. Data privacy protection is an important security requirement of the application service layer. Many information obtained by the IoT may contain personal privacy, such as positioning information obtained by GPS. Such information can be used by attackers to analyze users' sensitive privacy such as residence, income, lifestyle, behavior, and health status [20].

3) SOFTWARE SECURITY

Malicious applications are commonly used by software attackers. For example, in 2017, Bank of Russia found a malware called Bespalova existed in ATMs, which automatically paid after entering a specific code. If the system does not have enough code checks and tests, it will be vulnerable to attacks by malicious scripts or error indications. For example, attackers will use XSS (Cross-Site Scripting Attack) [21] to inject some malicious scripts into another trusted website. Successful XSS attacks can lead to hijacking IoT accounts and paralyzing the IoT system. In addition, Android malware has increased significantly in recent years [22]. For mobile devices, the openness of the Android mobile operating system has contributed to the spread of malware. Malware can use vulnerabilities to invade users' mobile phones and obtain private data.

III. FEASIBILITY ANALYSIS OF APPLYING

AI IN IoT SECURITY

A. COMMON CHARACTERISTICS AND SPECIAL REQUIREMENTS OF IoT SECURITY

Through the analysis of various security problems faced by IoT, it can be found that IoT security problems have some common characteristics which makes IoT security more complex and produces special requirements for security protection that are different from other fields.

1) THE DISTRIBUTION OF DATA REMAINS RELATIVELY STABLE IN NORMAL IoT CONTEXT, SO THE DETECTION OF ABNORMAL BEHAVIORS AND DATA OUTLIERS BECOMES THE MAIN IoT SECURITY REQUIREMENT

In general, due to the lack of resources, IoT devices can only complete simple tasks such as data collection and data transmission. A large number of common devices will maintain constant business modes and the collected data will maintain relatively stable distributions. For example, network traffic of consumer IoT devices generally has fixed modes. These devices tend to send stable signals to a limited number of endpoints, so their network activities are more predictable and structured [23]. On the contrary, DoS / DDoS attacks will generate significantly different network traffic than IoT devices. Therefore, real-time monitoring and identification of abnormal data, timely capture and early warning of abnormal business data flow are effective protection measures against many security attacks. Looking for technologies with the ability to distinguish normal and abnormal modes efficiently is the main requirement of IoT security.

2) THE UNPREDICTABILITY AND VARIABILITY OF ATTACK MODES LEAD TO THE LACK OF PRIOR KNOWLEDGE, WHICH PUTS FORWARD HIGHER REQUIREMENTS FOR THE ROBUSTNESS AND GENERALIZATION ABILITY OF SECURITY PROTECTION MODELS

In the past, the information security risk existed in the aspect of personal privacy. But with the expansion of IoT market

scale, IoT attack scenarios also show a trend of diversification. Attacks against hardware and software vulnerabilities, communication interfaces, or cloud platforms are emerging in endlessly. However, defenders lack of prior knowledge [24] about new attack modes and cannot adopt appropriate countermeasures in time. They can only understand the attack after bearing the loss, which greatly increases the security risk of users. If there is not enough prior knowledge, then security models need to be able to maintain the effectiveness in a variety of unknown scenarios, which requires higher robustness, generalization ability and data control ability of security protection models.

3) THE SECURITY MECHANISM OF LOW-POWER DEVICES AND MICROSERVICE TERMINALS LACKS THE ABILITY OF AUTONOMOUS PROTECTING, LEARNING AND UPGRADING

The large-scale interconnection of IoT devices requires low-power and low-cost solutions, so complex security mechanisms cannot be used. Therefore, current IoT security protection methods have the difficulty of updating outdated security strategies. Most of them do not have abilities of self-renewal and evolution, which is far from the "active immunity" or "auto-immunity [25]". Due to the lack of initiative, these methods rely on human to maintain and update the database, define new attack modes and interception rules. A lot of manual participation leads to a certain degree of lag in security protection, and it is unable to learn and upgrade the security scheme in time. How to design automatic security mechanisms with "active immunity" has become an increasingly urgent problem of IoT security.

4) THE INTEGRITY OF IoT REQUIRES THAT SECURITY SCHEMES CAN EFFECTIVELY HANDLE MASSIVE DATA AND CAN BE DEPLOYED IN A LARGE SCALE AND UNIFIED WAY

IoT is a multi-layer system across terminals, networks and service platforms, the defense of security issues is holistic and data are more complex and large-scale. The unity and integrity of IoT put forward an urgent demand for the processing capacity of security solutions under massive data, and it is necessary to ensure that solutions can be uniformly deployed in a large-scale way, remain effective in complex situations, and can evolve according to new scenarios at any time.

B. NEW CAPABILITIES PROVIDED BY AI TO IoT SECURITY

The particularity of IoT security and limitations of traditional methods highlight the urgent need for new security technologies. As a new technology direction, artificial intelligence has a wide range of applicability [26]. Machine learning (ML) is a research focus in the field of artificial intelligence. Its theory and methods have been widely used to solve complex problems in many engineering applications. The ML algorithms applied to IoT security can be divided into transaction algorithms and decision algorithms. (Fig. 2).

Transaction algorithms are mainly responsible for data exploration and data preprocessing. Transaction algorithms

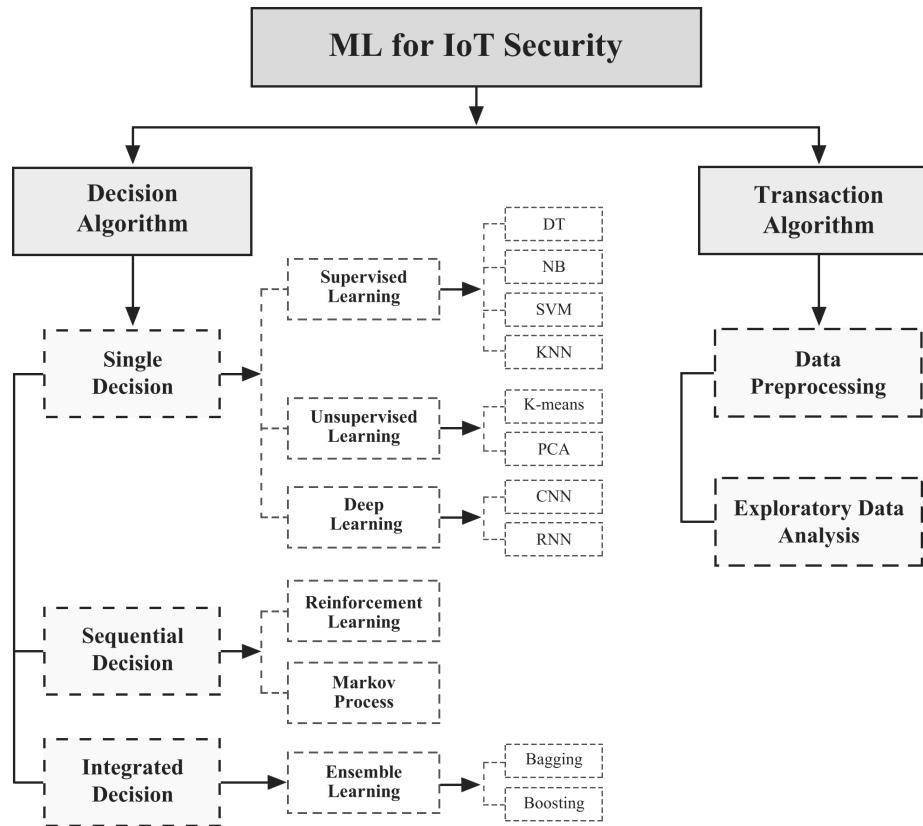


FIGURE 2. Machine learning for IoT security.

use few samples and simple models to obtain the general characteristics of the dataset and provide the basis for decision algorithms. Decision algorithms are mainly responsible for business decisions and adopt different decision-making strategies to reduce the ratio of misjudgment, so that the overall profit is the highest. Decision algorithms can be divided into three types according to strategies and scenarios: single decision-making, sequential decision-making, and integrated decision-making.

In addition, machine learning methods can be divided more carefully (as shown in Table 1), including Supervised Learning [27], Unsupervised Learning [28], Reinforcement Learning (RL) [29], Ensemble Learning [30], and Deep Learning (DL) [31].

Machine learning can make up for the defects of traditional security solutions in different aspects, conform to the characteristics of IoT and provide new capabilities for IoT to meet the new security requirements mentioned above.

1) THE CAPABILITY OF PATTERN RECOGNITION AND ABNORMAL BEHAVIOR DISCRIMINATION

Based on analysis of most IoT security incidents, we know that the business modes of IoT devices are fixed and their normal behaviors are predictable and structured. ML methods such as supervised learning and unsupervised learning can provide powerful ability to capture abnormal activities

and distinguish abnormal patterns, so as to classify normal behaviors and abnormal attacks. So, supervised learning and unsupervised learning can be widely used in IoT security.

For example, a specific example is to use Support Vector Machine (SVM) [32], a supervised learning method, to solve whether an access device is authorized. SVM can use a hyperplane to divide points in the feature space of device data into two categories: blue nodes are authorized devices, and yellow nodes are unauthorized devices, so as to classify different devices and intercept the unauthorized devices (as shown in Fig. 3).

2) THE CAPABILITY OF AUTONOMOUS PROTECTING, LEARNING AND UPGRADING

The lack of learning, upgrading abilities in unknown scenarios is one of the important reasons for the limited applicability of traditional schemes in practical applications. Traditional security schemes are not prepared enough for new viruses or attacks and cannot provide timely and effective means of resistance. For example, the important defect of Intrusion Detection based on misuse detection is the lack of capability of unknown network intrusions such as Zero Day attack [33].

AI provides automation and intelligence capabilities for IoT security in three aspects. First of all, unsupervised learning can automatically obtain knowledge from the data without known tags. The failure to obtain sample labels is a common

TABLE 1. Common ML methods for IoT security.

Type	Method	Advantages	Disadvantages	Common Application
Supervised Learning	Decision Tree [41]	High calculation efficiency; decision-making process is intuitive; no domain knowledge and parameter assumptions are required; suitable for high-dimensional data.	The structure of decision tree will be more complex when data with multiple features are involved, so appropriate pruning is needed to avoid over-fitting.	Device authentication; DoS/DDoS attack detection; Intrusion detection
	Naive Bayes [42]	The training process is simple, generally used for small-scale data sets; when the feature independence is satisfied, Naive Bayesian can also perform well in the efficiency of large-scale data sets.	The principle of feature independence is difficult to satisfy in many cases. When the assumption of feature independence fails, naive Bayesian is no longer applicable.	Device authentication; Intrusion detection
	SVM [32]	When data is not linearly separable, kernel method [43] can be used for nonlinear classification; SVM is especially suitable for datasets with a large number of features but a small number of samples; some support vector samples determine the final classification results, adding and deleting nonsupport vector samples have no impact on the model, so the robustness of SVM is good.	SVM is difficult to implement for large-scale datasets. When the number of samples is large, the storage and calculation of matrix will consume a lot of memory and time. And SVM is difficult to solve multi-classification problems.	DoS/DDoS attack detection; Intrusion detection; Malware detection
	KNN [44]	KNN is simple and easy to understand. The calculation method is still valid for large-scale datasets. KNN can be used for nonlinear classification and multi-classification problems.	It is a time-consuming process to determine the best value of K.	DoS/DDoS attack detection; Intrusion detection
Unsupervised Learning	K-Means [34]	K-Means can be applied to unlabeled samples to learn the representation of input data without prelabeled training data.	It cannot deal with nonspherical clustering; the selection of optimal K is a problem.	DoS/DDoS attack detection; Intrusion detection
	PCA [45]	The calculation of PCA is simple and has no parameter limitation.	Principal component is not as explanatory as the features of the original samples; when the variables do not obey the Gaussian distribution, scaling and rotation will occur.	Feature selection or Feature reduction
Ensemble Learning	Random Forest [46]	The effect of Random Forest is better than single models; the introduction of randomness makes RF not easy to overfit, has a good anti-noise ability, and is not sensitive to outliers; it can handle high-dimensional data; it has no requirements for data types, and can handle both discrete data and continuous data.	The features with more value division will have greater impact on random forest; in some classification or regression problems with more noise, they may be over-fitting.	Device authentication; DoS/DDoS attack detection; Intrusion detection; Malware detection
Reinforcement Learning	Q-Learning [47]	Q-Learning requires few parameters and can be implemented offline.	The search and storage of Q-table need a lot of time and space; update speed is slow; not suitable for high-dimensional data.	Malware Detection
Deep Learning	CNN [48]	CNN shows applicability and robustness in many image applications and is an effective image classification and recognition model.	DL needs a lot of data, computing resources and high hardware requirements; black box model, poor interpretability; complex model structure, high computational complexity.	Device authentication; DoS/DDoS attack detection; Intrusion detection; Malware detection
	RNN [49]	RNN have achieved excellent performance in machine translation, speech recognition and other applications with sequential data.		

lack of prior knowledge, unsupervised learning extends from relying on label data to using unlabeled data, which can effectively reduce the consumption of manually labeled data and maintain its effectiveness in the scenario without empirical data. For example, unsupervised clustering methods, represented by K-Means [34], can divide the input data into different clusters without labels by detecting the similarity between input data. As shown in Fig. 4, when $K = 3$, the K-Means algorithm divides all the sample points into three clusters, and then determines the node risk of each cluster according to the common properties of the samples in each cluster. Then, the security level of IoT nodes can be divided so as to take different countermeasures.

Secondly, ensemble learning can make use of the results of multiple classifiers to vote and adjust the learning focus of the model, so as to gradually and automatically improve the model effect and avoid multiple manual operations. The aggregation of models from multiple scenarios can also effectively reduce the deviation caused by a single scene and enhance the applicability to new scenarios.

In addition, reinforcement learning can achieve the gradual optimization of the model through the reward / punishment mechanism and adjust learning strategies in a constantly changing environment to maximize the benefits. RL can learn new strategies while constantly inputting new data, ensure security models of IoT adapt to the changes of new

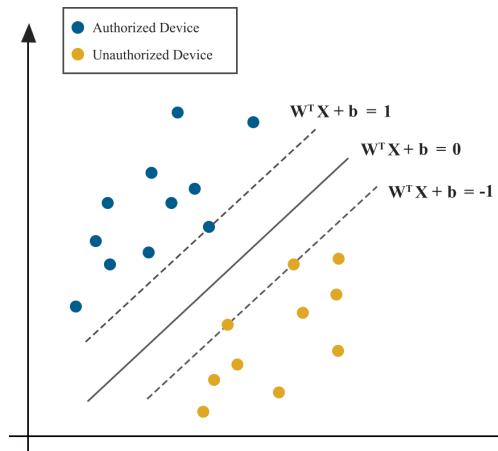


FIGURE 3. An example of IoT devices classification based on SVM.

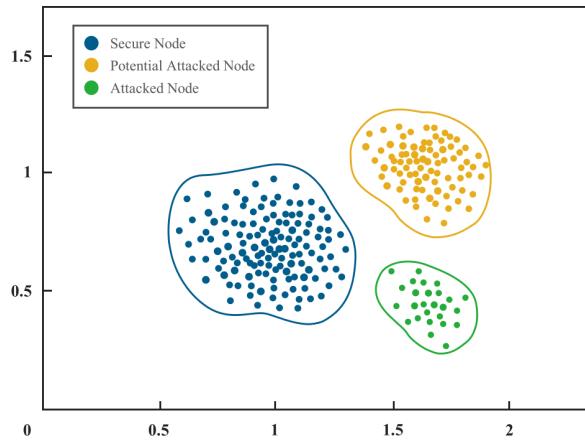


FIGURE 4. An example of IoT node clustering based on K-Means ($K = 3$).

environment, maintain the validity, and strengthen the active exploration ability of the model, thus laying the foundation for the realization of active immunity of IoT security.

3) THE CAPABILITY OF PROCESSING LARGE AMOUNTS OF COMPLEX DATA EFFECTIVELY

Traditional IoT security schemes usually work under the limited amount of data. With the increasing generation of data, the deficiencies of these schemes in big data processing capacity and computing efficiency are highlighted. For example, malware detection is the primary task of software security in application service layer. Traditional malware detection methods extract malicious behavior codes from malware as signatures, and judge whether a new software is malware by calculating the similarity between the software to be detected and the signature database. When the amount and dimensions of data increase, the computational complexity will rise rapidly, resulting in the efficiency of the model is greatly reduced, and it is unable to make timely and effective detection.

Compared with traditional schemes, the advantage of AI schemes is that it can not only process small-scale data,

but also can make use the data set with more samples and higher dimensions for effective calculation. For example, ImageNet [35] is one of the most famous data sets in the field of image processing. Its data volume has reached 10 million, and many of its sub datasets also have a million data. Even so, many deep learning models have achieved very good results on some ImageNet sub datasets.

4) THE CAPABILITY OF MAINTAINING EXCELLENT MODEL ACCURACY

The low model accuracy of traditional models is also the reason why AI technologies are urgently needed in IoT security. If models with poor effect cannot detect the potential attack in time, it will reduce the credibility of security models. In addition, due to technical limitations, some traditional models may damage the operation mechanism of IoT in order to improve the model effect. For example, SYN Flood attack [36], a very common DDoS attack, sends a large number of attack messages with forged source address to the network service port, causing the connection queue in the target server to be occupied. A protection measure is called random drop (RD). By randomly discarding SYN requests in TCP backlog queue, RD can alleviate the queue pressure and reduce server load. However, the success rate of many legitimate clients with normal connection or slow connection will be greatly reduced, even unable to respond to SYN-ACK messages from the server, thus destroying the operation mechanism of TCP itself [37].

AI schemes can effectively calculate large data sets, at the same time, it can ensure that models have good effectiveness and perform well in many evaluation indicators such as precision and recall. Many supervised learning methods and deep neural networks have provided excellent solutions in various application scenarios.

5) THE CAPABILITY OF PROVIDING MODEL ROBUSTNESS AND GENERALIZATION ABILITY

Traditional security solutions generally solve problems in relatively simple environments. When these solutions migrate to more complex scenarios, such as enterprise IoT security, they may not achieve the expected results. For example, the traditional password device authentication, due to the lack of complexity of the password form, is easy to lead to password impersonation and interception, so password device authentication is only applicable to closed small systems [60].

AI methods pay great attention to the robustness and generalization ability. Robustness [38] requires the model can effectively reduce the impact of noise and outliers and ensure the model to maintain effectiveness in complex scenes. The generalization ability [39] reflects the prediction ability of the model for unknown data, which ensures that the model will not lose efficacy after being transferred from the experimental scenario to the application scenario.

A variety of ML methods have advantages in robustness and generalization ability: SVM determines the classification result through a small number of support vector samples,

adding or deleting nonsupport vector samples has no effect on the model, which makes SVM have good robustness; random forest has good anti-noise ability and is insensitive to outliers; linear models with L1 and L2 regularization [40] has excellent generalization ability and can avoid over-fitting. These ML methods with good robustness and generalization ability can greatly enhance the applicability and scalability of IoT security solutions.

IV. AI SOLUTIONS TO FOUR IoT SECURITY THREATS

A. FOUR THREATS TO IoT SECURITY

According to investigation, there are four threats that need to be solved urgently in IoT security: device authentication, DoS/DDoS attack, intrusion detection, and malware detection. The traditional solutions to these problems lack the processing ability of large data sets and have many problems such as low efficiency and poor real-time performance. Most of them cannot be migrated to IoT. AI methods represented by machine learning can use massive IoT data to infer useful knowledge from the data, and thus make predictions for unknown events, providing new solutions for these problems.

1) DEVICE AUTHENTICATION

There are risks of interception, counterfeiting, tampering, and destruction in the process of information interaction and data transmission between IoT nodes. In order to prevent the transmission of false information, the security requirements between nodes include identity authentication, judgment and blocking malicious nodes [50]. The authentication process of IoT devices is generally restricted by the characteristics of the IoT, such as limited resources. Therefore, it is necessary to ensure that the calculation and communication cost do not exceed the limitations of the device as much as possible, to ensure that the device does not consume too many resources [50].

2) DoS / DDoS ATTACK

Denial-of-service attacks (DoS) [51] and distributed denial-of-service attacks (DDoS) [52] use weaknesses in the transmission protocol, or vulnerabilities in systems and servers to launch large-scale destructive attacks on the target system. Massive data packets exceeding the target processing capacity will consume available network bandwidth resources, causing program buffer overflow, preventing other legitimate users from normal requests, and ultimately lead to network service paralysis or system crash. There are some differences between DDoS and DoS. DDoS uses multiple distributed attackers in different positions to launch attacks on one or several targets at the same time, or an attacker controls multiple machines in different positions and uses these machines to attack the victim.

3) INTRUSION DETECTION [53]

Intrusion Detection aims to monitor events that occur in the system, by collecting and analyzing the information of key

points, to check whether there are behaviors that violate the security policy to achieve early detection of intrusion. As an active security protection technology, intrusion detection is an important part of network security providing real-time protection against internal and external attacks. The ability to detect intrusions and malicious activities in the IoT network is critical to the timely recovery of network infrastructure.

4) MALWARE DETECTION [54]

IoT allows a large number of smart devices to connect to each other to share information and improve the user experience. In order to provide interactive services with users, more and more PC or mobile applications appear. Using vulnerabilities of these applications to inject and execute malicious code in IoT software is a common attack method. These vulnerabilities that can be used for malware injection may be related to the authentication and authorization of the application. Physically tampering with IoT devices, software modifications and misconfiguration of security parameters may also allow attackers to inject malicious code. Common malware includes bots, ransomware, adware, etc.

B. GENERAL PROCESS OF AI SOLUTIONS FOR IoT SECURITY

The main tasks of device authentication, DoS / DDoS attack detection, intrusion detection, and malware detection are classification tasks. For example, for device authentication, AI solutions need to be able to accurately classify authorized and unauthorized devices; for intrusion detection, the solutions need to be able to classify normal and abnormal network behaviors; and so on. We analyze existing machine learning solutions to these problems and summarize the flow of most solutions into the basic process shown in Fig. 5.

1) DATA COLLECTION

ML solutions usually require data sets from specific environments. For different problems, you need to choose the appropriate environment for data collection to form training data sets and test data sets. For example, data sets of device authentication need to contain the information of device configurations, user behavior and operation habits to reflect the differences of users.

2) DATA PRE-EXPLORATION AND PRE-PROCESSING

The quality of training data is directly linked to the effect of solutions. IoT data sets comes from various sensors in various fields. However, there are more or less problems in the original data set, such as irregular data distribution and incomplete data. Therefore, it is necessary to mine the training data, master the data distribution, and then carry out operations such as deleting errors and completing incomplete data, so as to prepare for subsequent steps.

3) MODEL SELECTION

There are many ML models that can be selected for IoT security, but each model has its own applicable scenario,

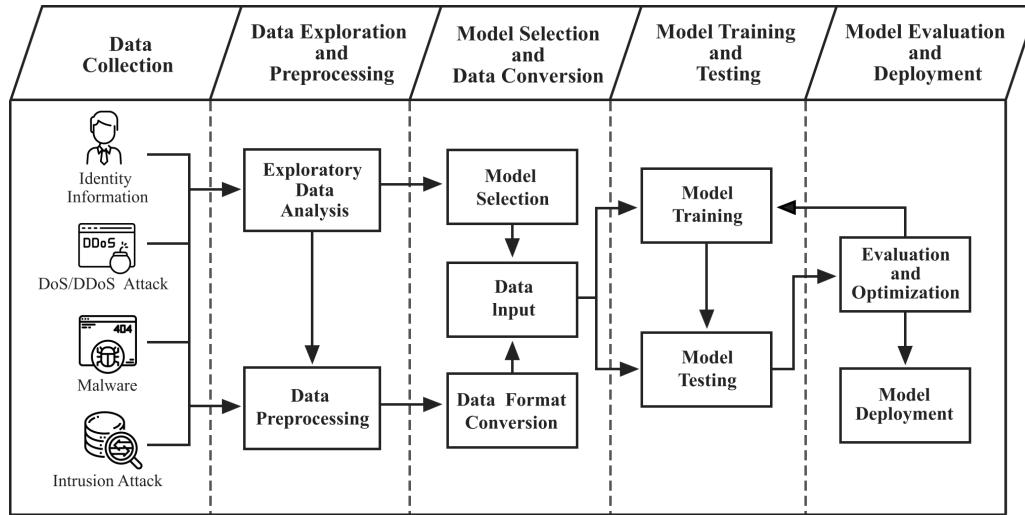


FIGURE 5. The general process of AI solutions for IoT security.

so we should select appropriate models according to the characteristics of models and requirements of problems. The size of the data set and the pre-exploration results of the data will also affect the choice of the model. For example, if the data set has fewer simple samples and the training needs to be completed in a short time, then the lightweight algorithms such as naive Bayes can get expected results.

4) DATA CONVERSION

In actual applications, the data collected is usually inconsistent with the input data required by models and needs to be converted to meet the needs of models we selected. For example, the audio data collected by voice sensors cannot be directly input into RNN models, so it is necessary to carry out data conversion. One of the conversion methods is to extract the Mel-Frequency Cepstral Coefficients (MFCC) from original audio data [55].

5) TRAINING AND TESTING

After data pre-processing and model selection are completed, we need to input data into models for training. In the training process, we can observe the loss function value or result curve to know the training trend of the model, so as to adjust learning rate or other parameters appropriately to ensure that model effect is gradually optimized. After obtaining training models, test datasets obtained from the real world are used to test the generalization ability of training models. Training models may be under-fitting or over-fitting [56], so parameters need to be adjusted again.

6) MODEL EVALUATION AND DEPLOYMENT

When selecting the final model for actual deployment and application, we can use some effect indicators to evaluate different models after training. Different evaluation indicators are selected according to problems such as classification,

regression, ranking, to objectively evaluate the prediction and generalization ability of the model. For IoT security, commonly used evaluation indicators are accuracy, precision, recall, F1 score [57], AUC (Area Under ROC Curve) [58], etc.

C. AI FOR DEVICE AUTHENTICATION

1) TRADITIONAL DEVICE AUTHENTICATION: DIGITAL CERTIFICATE AUTHENTICATION AND PASSWORD AUTHENTICATION

Digital certificates include user identity information, relying on a trusted third-party Certification Authority (CA) to achieve authentication, and users can access servers of CA with the authentication certificate. The International Telecommunication Union's X.509 standard defines a framework for providing authentication services. General CA digital certificates follow X. 509 standard format, so it is also called X. 509 certificate [59]. The password authentication saves the user's name and password in advance [60]. When the user enters the system, the entered information is compared with the previously saved information to verify whether the user's identity is legal.

Traditional authentication technologies have many problems. For example, although the password authentication is simple and easy to implement, it is generally only suitable for closed systems. Every time users access the system, they must enter the password in plaintext, which may be intercepted in the process of transmission, thus revealing the privacy information. Traditional IoT terminals are easy to be counterfeited because of the static nature of identification information [61]. The static of device ID or user ID makes the identity easy to be scanned, read and counterfeited by hackers. ML provide a variety of feasible ideas for secure authentication of IoT. These schemes use a variety of ways to obtain verifiable information related to devices and users. We select several representative schemes for comparative analysis (as shown in Table 2).

TABLE 2. Comparison of representative device authentication schemes.

Scheme	Information Used in Scheme	Technology	ML Methods Used in Scheme
Traditional methods	Third-party digital certificate; user name and password in plaintext	Digital certificate authentication; password authentication	No ML methods used
[62]	Network traffic	White lists	Ensemble Learning
[63]	Acoustic characteristics of human respiration	Breath Print	RNN
[65]	Channel Status Information (CSI)	Wi-Fi	DNN
[61]	Operation behaviors and device information	Dynamic device fingerprint	Decision Tree; Naive Bayes; Logistic Regression
[66]	Communication signals	RF wireless communication fingerprint	PCA; Random Forest; ANN

2) MANUALLY SET INFORMATION: WHITE LIST OR BLACK LIST

The white list and black list are manually set safety / dangerous device lists, which are often used to screen and intercept connected devices. The goal of device authentication is to ensure that only authorized IoT devices can connect to the network, but there are too many and increasing devices with vulnerabilities, which makes organizations be cautious and skeptical about all IoT devices. A white list of authorized devices will be much smaller than a black list of increasing potentially dangerous devices and can also improve the efficiency of ML model training, testing, and deployment. Meidan *et al.* [62] designed an authentication scheme combining white list and ensemble learning, using random forest to perform feature extraction and devices classification of network traffic data in large enterprises IoT. For testing, the average accuracy of the nine device types tested was 99%, and the method was desirable in accuracy and speed of classification.

3) HUMAN BIOLOGICAL CHARACTERISTICS

Human biological characteristics refer to the inherent characteristics of the human body such as fingerprints, irises, faces, DNAs, sounds, and so on. Sound sensors in wearable IoT devices such as smartphones and watches can be used for identity authentication. These devices often interact with individuals frequently to obtain personal-specific information and have unique advantages in fine-grained monitoring of user environments. Breath Print is an authentication technology for respiratory acoustics on mobile IoT devices. Breath Print assumes that each person's breathing pattern is unique, thereby taking advantage of the user's respiratory acoustic characteristics captured by wearable IoT devices to support user authentication. With the unique advantages of RNN in audio and speech processing, Chauhan *et al.* [63] combined RNN with Breath Print to model the collected respiratory acoustic data to distinguish different users. Experiments showed that this method can be effectively implemented on various resource-constrained embedded devices with low latency (less than 200ms for smartphones). It should

be noted that human biometrics are important privacy data for users, so it is necessary to prevent possible privacy disclosure when using them. The Cancellable Biometric System (CBS) technology can convert the forms of original biometric data. Then, the transformed data can replace the original biological characteristics at any time, so as to protect the original data from being destroyed. Punithavathi *et al.* [64] developed a safe prototype of lightweight cancelable biometric recognition system with the help of image preprocessing, feature extraction, feature conversion, template matching and other machine learning technologies, which solved the privacy problem of using human biological characteristics.

4) HUMAN BEHAVIOR CHARACTERISTICS

Human behavior characteristics such as gaits, handwritings, object manipulation habits are also commonly used identity authentication information. Electronic devices in the indoor environment (such as smart refrigerators, smart TVs, smart air conditioners, and security doors) can obtain human behavior characteristics. There are rich Wi-Fi signals between these devices. When operating these devices (such as opening refrigerator doors, entering or leaving room), it is possible to capture the unique physiological and behavioral characteristics of human daily activities, providing a feasible direction for distinguishing each individual. Recognizing user activities needs to start from simple actions and rise to the unique behaviors of different users. The system needs to have abstract capabilities with different granularities, which can extract different levels of feature representation. The powerful abstract representation capabilities of deep learning provide the possibility for this. Shi *et al.* [65] used the amplitude and relative phase of the Channel State Information (CSI) in Wi-Fi signals of household appliances to extract representative human behavior features, combined with the three-hidden-layer DNN model to abstract the features at different levels (Fig. 6). This method achieved the authentication accuracy of 94% and 91% for dynamic and static human activity identification, which proved the feasibility of the combination of Wi-Fi signals and DL.

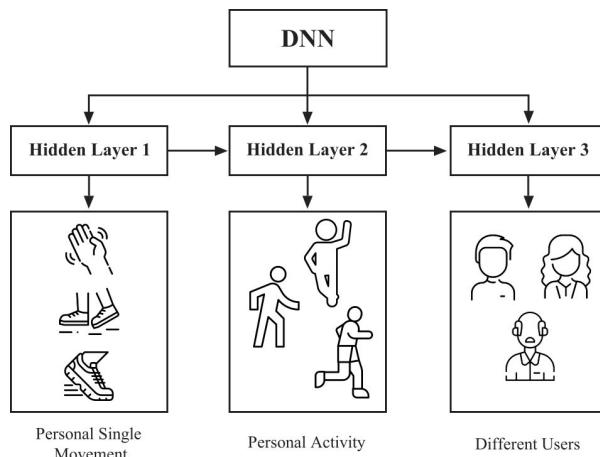


FIGURE 6. Abstract representation capability of DNN with three hidden layers.

5) STATIC DEVICE INFORMATION

The RF fingerprints extracted from RF wireless communication devices can reflect the hardware differences of each device. Although the differences in hardware are small, it has been proved that these differences can be used for device authentication. Lin *et al.* [66] obtained RF fingerprints of IoT devices, and then used PCA to reduce the dimension of fingerprint features. The high dimensional fingerprint was reduced from 3187 dimensions to 2 / 76 / 300 / 645. Then, random forest, SVM and artificial neural network were used to identify the low dimensional data. Experiments showed that the detection accuracy of 76-dimension data was better than that of other dimensions and random forest performed better than other algorithms.

6) DYNAMIC OPERATING BEHAVIOR INFORMATION

Dynamic terminal fingerprint is the combination of static device information and user dynamic operation behaviors. It can contain information such as IP address, operating system version, port status and network access location, etc. Through ML methods, the difficulty of authentication and recognition caused by the dynamics of device fingerprint can be solved. Zhang *et al.* [61] proposed an intelligent identification scheme combining dynamic terminal fingerprints with decision trees, logistic regression and naive Bayes. The classification accuracy based on decision tree reached 98% and performed better than logistic regression and naive Bayes. Bezawada *et al.* [67] used the protocol list used by IoT devices in various stages of their operation (such as ARP, HTTP, DNS, etc.) as static behavior information, and used the features extracted from the network traffic of the devices as dynamic behavior information. The fingerprint composed of static and dynamic behavior information can combine with ML models such as KNN, decision tree and gradient boosting. In their experiments, the average accuracy of detecting similar types of devices was up to 99%.

D. AI FOR DoS/DDoS ATTACK DETECTION AND DEFENSE

The cleaning and filtering of abnormal traffic is the key of DoS / DDoS protection. Traditional DoS and DDoS defenses are optimized for load balancing, and use anti-DDoS devices, firewalls, or other protection settings. These methods judge whether the external access traffic is normal through firewall, rule filtering and content filtering. However, due to the large number of devices and limited resources in the field of IoT, it is more and more difficult for IoT devices to resist DoS and DDoS attacks, which requires efficient and accurate traffic filtering methods.

Machine learning is a good choice for providing intelligent and automated DoS / DDoS detection mechanisms. DoS and DDoS attacks in a variety of different IoT scenarios can be solved through machine learning (as shown in Table 3).

1) SOFTWARE-DEFINED NETWORK

Software-defined

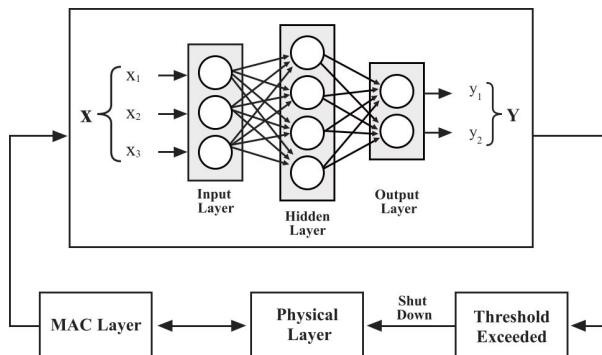
network (SDN) is a network architecture that separates the control plane and data plane of network devices. The rich functions provided by the SDN control plane enable organizations to effectively control IoT devices. With the characteristics of centralized control, flexibility, and scalability, SDN can be used as the underlying communication infrastructure of IoT. However, the openness of the interfaces in SDN also brings security implications and is vulnerable to DDoS attacks. Ye *et al.* [68] proposed a DDoS detection scheme based on SVM, which considered attack detection as a classification problem. The scheme extracts the values in the switch flow table of the SDN architecture related to DDoS attacks, such as the speed of source IP, source port and flow entries, the standard deviation of flow packets, the deviation of flow bytes, the ratio of pair-flow, to use as characteristic values for SVM classification.

2) WIRELESS SENSOR NETWORK

Wireless sensor network (WSN), as an important communication technology in IoT system, also has the danger of being targeted by DoS attacks. Designing a secure media access control (MAC) protocol against WSN-oriented DoS attacks is currently an important research direction, and deep learning has provided assistance for this. Kulkarni *et al.* [69] proposed a new secure MAC protocol based on Multilayer Perceptron (MLP) (Fig. 7). In this MAC protocol, each node in WSN has a trained MLP running on the MAC layer. This new MAC protocol extracts key parameters from the environment as MLP's inputs: collision rate, packet request rate, average packet waiting time and so on. MLP will output the probability (suspicion factor) of nodes suffering DoS attack. If the suspicion factor is greater than the predefined threshold, the node will shut itself down and save energy within a preset duration, limiting the attack to a local area of the network. At the same time, shutting down the attacked nodes also reduces the power consumption and prolongs the service life of the system.

TABLE 3. Comparison of representative DoS / DDoS detection schemes.

Application Scenario	Information Used in Scheme	Method Characteristics	ML Methods Used in Scheme
Software-defined Network [68]	Speed of source IP; Speed of source port; Standard deviation of flow packets; Deviation of flow bytes; Speed of flow entries; Ratio of pair-flow	This method uses an adaptive and accurate classifier to make decisions based on uncertain information, which can detect attacks early.	SVM
Wireless Sensor Network [69]	Collision rate; packet request rate; average waiting time of packets	This method can limit the scope of DoS attack, close the attacked nodes in time, reduce the consumption and extend the service life of system.	MLP
Consumer IoT [23]	Network flow characteristics	The method shows that the home gateway router can use low-cost ML algorithms and traffic data to automatically detect the DDoS attack source of local IoT devices.	KNN; Decision Tree; Neural Networks; Random Forest; SVM
Smart City [70]	41 features such as duration, protocol, service, flag, source bytes, destination bytes, etc.	It has autonomy and effectiveness of local attack detection, reduces storage and computing consumption, and provides fast response.	Deep Learning

**FIGURE 7.** Secure MAC protocol combined with MLP.

3) CONSUMER IoT

Consumer IoT are also at risk of DDoS attacks, such as wearable devices and smart appliances. The IoT traffic of these devices is different from attacks. ML can capture traffic that differs greatly from the network traffic characteristics of specific behaviors of consumer IoT devices, such as DDoS traffic, for attack detection. Doshi *et al.* [23] used a variety of ML algorithms such as K-nearest neighbors, decision trees, neural networks, random forests and SVM to achieve high-precision DDoS detection in consumer IoT traffic. The results showed that the home gateway router can use low-cost ML algorithms and traffic data to automatically detect the DDoS attack source of local IoT devices.

4) SMART CITY AND SOCIAL IoT

Smart city is the fastest growing and most influential public service field of social IoT, which helps the city effectively manage water, electricity, transportation and other infrastructure. Millions of users are connected to social

services through IoT, taking advantage of the private and public services provided by IoT. DoS attack has become the most frequent attack type in smart city. Diro *et al.* [70] used deep learning and fog computing architecture to train models and hosted attack detection systems at the edge of the distributed fog network. The combination of deep learning and fog operation enhanced the autonomy and effectiveness of attack detection in local model, accelerated the speed of data training and reduced the calculation costs of model, data, and parameters of IoT.

E. AI FOR INTRUSION DETECTION

1) TRADITIONAL INTRUSION DETECTION: MISUSE DETECTION OR ANOMALY DETECTION

With the development of IoT, new types of intrusions have brought new problems to intrusion detection. Intrusion through unauthorized access to obtain confidential information is still increasing, but technologies such as access control, data encryption, and firewalls have certain limitations. Most current intrusion detection systems use misuse detection or anomaly detection. Misuse detection [71] can effectively detect known attacks according to attack patterns that have appeared. However, they cannot detect unknown new types of intrusions such as zero-day attacks, because these attacks are not similar to known attacks. In contrast, anomaly detection [72] analyzes normal traffic patterns and makes judgments based on the assumption that attacker's behaviors are different from normal users. If the characteristics of a certain traffic are far from normal traffic patterns, the traffic is considered an intrusion. Anomaly detection is useful for new attacks, but they are not as effective as misuse detection in terms of known attacks. Identifying anomalous activity from massive data is a time-consuming process.

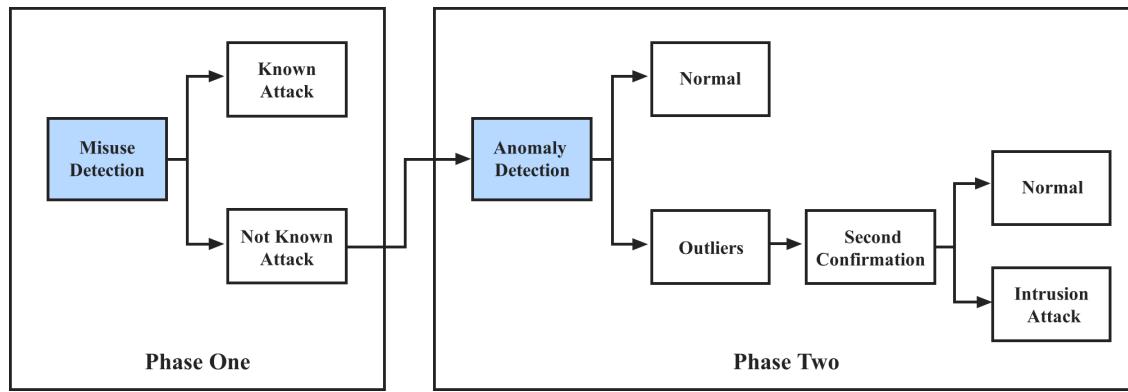


FIGURE 8. Combined intrusion detection.

Anomaly detection still needs to be developed in terms of shortening time and improving accuracy.

ML have been applied to intrusion detection and have shown better performance than traditional methods. ML have brought new changes to intrusion detection and have produced several new development trends: from single models to combined models, from focusing on method accuracy to taking into account both efficiency and effectiveness, from centralized detection to decentralized detection.

2) FROM SINGLE MODELS TO COMBINED MODELS

In order to solve the shortcomings of two above-mentioned detection methods, some studies have proposed hybrid intrusion detection that combines misuse detection and anomaly detection. Most hybrid detection systems independently train misuse and anomaly detection models, and then aggregate the results of two models. Unlike usual methods, Kim *et al.* [73] proposed a new combined detection method that integrates a misuse detection model and an anomaly detection model hierarchically (Fig. 8). In the process of integration, misuse model can capture known attacks, and then use anomaly model to supplement the ability of unknown attack detection. The program first used training data composed of normal traffic and known attack traffic to train a C4.5 decision tree to build a misuse detection model, and then trained 1-class SVMs on unknown attack sub dataset of the C4.5 decision tree branches to establish multiple anomaly detection models. Through the test, the combined model performed better than single traditional model in detecting both unknown attacks and known attacks.

3) FROM FOCUSING ONLY ON ACCURACY TO TAKING BOTH EFFICIENCY AND EFFECTIVENESS INTO ACCOUNT

Generally, most intrusion detection systems emphasize effectiveness and ignore efficiency. For intrusion detection systems, too many data features may not guarantee good performance, but will increase decision delay, so it is critical to choose fewer but more important features. The feature selection of ML can be a method to improve the efficiency of intrusion detection.

Li *et al.* [74] designed a two-stage combined model which takes both efficiency and effectiveness into account. In the first stage, they used the heuristic iterative search ability of Swarm Intelligence (SI) algorithm to search for the optimal features. In the second stage, they used the features selected in the first stage as inputs and used random forest to classify the network traffic into different attack categories. The model selected more important features, improved the operation efficiency of the model, and achieved better performance in intrusion detection. Su [75] proposed a feature selection scheme combining KNN and GA to improve the efficiency of intrusion detection. A major disadvantage of KNN is that each data feature is given the same weight, but in fact some features may be more important than others, and some features can be ignored. The goal of GA combined with KNN is to find an optimal feature weight vector. Genetic algorithm (GA) is a search algorithm to find the optimal solution by simulating the natural evolution process. GA starts with a population that represents the potential solution set of the problem. According to the principles of “survival of the fittest”, successive generations produce better and better approximate solutions. After the evolution of GA, the optimal weight vector of KNN can be obtained. The researchers used the header information of network protocol including IP, TCP, UDP, ICMP, ARP, and IGMP to extract 35 features as initial training data, used the proposed algorithm to select some important features for intrusion detection. For known attacks, when considering the first 19 features, the overall accuracy was 97.42%. For unknown attacks, the accuracy of using the first 28 features was 78%. The scheme greatly improved the time efficiency of intrusion detection under the premise of basically guaranteeing the effectiveness of detection.

In addition to feature selection methods can improve efficiency, the above-mentioned combined intrusion detection [73] can also reduce computational complexity of models by decomposing dataset. When training dataset is decomposed into smaller subsets, the training and testing time will be significantly reduced to achieve the purpose of improving time efficiency.

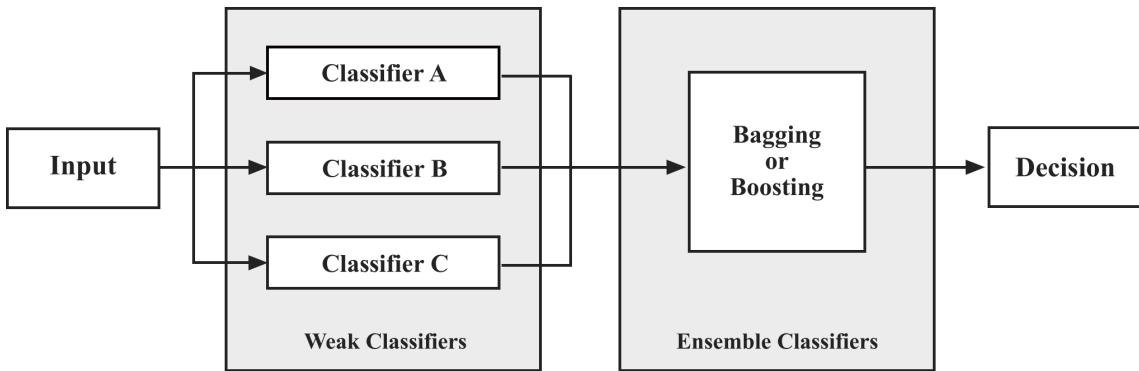


FIGURE 9. Enhancing weak classifiers by ensemble learning.

4) FROM CENTRALIZED DETECTION TO DECENTRALIZED DETECTION

At present, some ML or data mining intrusion detection methods usually require large amounts of stored data and strong computing power to find intrusion attack patterns, so they need to be executed in large central computing systems. But these methods are incompatible with IoT. The number of IoT devices is increasing and most IoT devices have limited computing resources, but there will be no suitable centralized processing equipment. It is difficult for general devices to replace central computing nodes to complete their centralized computing tasks. For these reasons, the demand for decentralized intrusion detection is increasing.

Bosman *et al.* [76] proposed a decentralized anomaly detection framework that includes a set of heterogeneous local anomaly detection models. The framework first uses incremental learning to construct decentralized models with little or no prior information. The computational cost and memory cost of decentralized incremental learning models are usually low, which can meet the requirements of limited resources. In addition, since the detection accuracy is limited by the quality of models, the accuracy of single model may not reach the expected goal, so ensemble learning can be used as a suitable option to increase accuracy. By integrating the outputs of multiple decentralized weak classifiers, the accuracy of detection can be improved. As shown in Fig. 9, firstly, decentralized weak classifier A/B/C which meets the constraints of the embedded platform will be trained. Then multiple weak classifiers are enhanced by ensemble learning. Compared with centralized detection, decentralized method has proved that under the strict constraints of the embedded system, decentralized method is feasible. In environments with less prior knowledge, it eliminates the need for manual intervention and performs as well as centralized solutions.

F. AI FOR MALWARE DETECTION

1) TRADITIONAL SOLUTIONS: SIGNATURE-BASED MALWARE DETECTION

Signature-based malware detection [77] generates unique signatures for each known malware to create malware behavior libraries. These signatures are manually found by experts

or generated by automatic methods, and can include many information, such as file names, content strings or bytes. The signature of unknown software can be compared with the malware behavior library to search whether there are matching signatures. This method is the most convenient and widely used detection method with fast detection speed and low false alarm rate. But the same as the misuse intrusion detection mentioned above, the signature-based malware detection is powerless for the malware that has not appeared before. The malware library needs to be constantly updated and maintained. However, there must be an initial victim reporting malicious activity before any form of detection or prevention can be carried out. When the initial victim is important, the consequences may be unacceptable. For example, the critical infrastructure vulnerabilities of U.S. Office of Personnel Management in 2015 may lead to decades of chain reaction and events [78].

Various ML methods have been applied to malware detection. (as shown in Table 5) These ML schemes choose to decompose the malware, trying to extract the potential information of the software, so as to pave the way for the models to detect malware. How to choose appropriate input information becomes the key to AI malware detection. We analyze several representative ML malware detection schemes and compare several different potential software information that researchers choose to use.

2) NETWORK BEHAVIORS

The wireless multimedia system (WMS) is used to continuously collect information and control the status of remote devices. Wireless multimedia devices are usually equipped with multiple sensors, which transmit data to neighbors based on routing tables. The distributed topology of WMS accelerates the spread of malware, thereby threatening other nodes, wireless routers and terminals through data transmission. For the detection of WMS malware, obtaining its network behaviors is critical. The malware detection scheme of Zhou *et al.* [79] facilitated the collection of network behaviors in WMS using the data sniffer (DroidSniffer), combined with SVM, BP neural networks to detect malware and suppress malicious codes. In the experiment, the highest infection rate

TABLE 4. Comparison of representative intrusion detection schemes.

Scheme	Information Used in Scheme	Method Characteristics	ML Methods Used in Scheme
Misuse detection and Anomaly detection	Network traffic	Misuse detection can detect known attacks effectively but cannot detect unknown new attacks. Anomaly detection can detect new attack modes, but the detection rate of known attacks is lower than misuse detection	No ML methods used
Combined Model [73]	Network traffic	This method combines the advantages of misuse detection and anomaly detection to make up for the defects of single model. The structure of hierarchical integration is more flexible, and the efficiency can also be improved by decomposing data sets.	SVM; Decision Tree
Feature Selection [75]	The network protocol header of IP, TCP, UDP, etc.	This method can select important features while ensuring the effectiveness of the model without increasing the efficiency through feature selection.	KNN; Genetic Algorithm
Decentralized Detection [76]	Time series data of sensors	This method disperses the detection process to each independent node, which can be executed in the environment with limited computing resources. The detection node can improve the effect of multiple weak models through ensemble learning.	Incremental Learning; Ensemble Learning

TABLE 5. Comparison of representative malware detection schemes.

Scheme	Software Information Used in Scheme	Malware Type	ML Methods Used in Scheme
Signature-based detection [79]	A signature containing software information generated manually or automatically	No restrictions	No ML methods used
	Network Behaviors	WMS malware	SVM; BP Neural Network
[80]	APK; API	Android malware	Ensemble Learning
[82]	Binary image	Telnet attack software	CNN
[83]	Opcode	Windows malware	CNN

was only 22.17%, which proved that malware can be detected at a lower infection rate.

3) APK AND API

Android platform plays a crucial role for the rapid development of IoT applications. At the same time, malware for Android has also increased, and virus strains that use highly sophisticated evasion techniques have emerged. Yerima *et al.* [80] developed a high-precision Android malware detection solution based on ensemble learning. The key to the analysis of Android malware is the software features obtained from the APK. Java-based APK analysis tools are used to extract the storage of features from the app corpus. The extracted 65 features include various API calls and Linux/Android command sets. These APIs include: SMS manager API (used to send, receive, read SMS messages, etc.); Phone Manager API (used to access device ID, subscriber ID, network operator, SIM serial number, etc.); package management API (used to list installed packages). Similarly, Zhu *et al.* [81] pointed out that the extraction of sensitive data stream in the application can effectively detect malware. They built DeepFlow, an Android malware

detection tool, analyzed Android API codes with APK, extracted the sensitive data stream, and used Deep Belief Networks (DBN) for classification. Experiments on 3000 benign applications and 8000 malicious applications showed that DeepFlow achieved a high F1 score of 95.05% with appropriate parameters.

4) BINARY IMAGE

A novel malware detection method is to analyze the binary image transformed from software. The binary file of software can be reformatted into an 8-bit sequence and then converted into a grayscale image, which has one channel and pixel values from 0 to 255. The experiment of [82] gives converted images and points out that the structural difference between the benign software image and the malignant software image is obvious. The malware image is always denser, for example, most Mirai malware images have dense centers. Su *et al.* [82] used the difference in binary images of different software to transform malware detection into an image classification problem, so as to distinguish benign software and malware by convolutional neural network (CNN). The experimental results showed that the classification accuracy of CNN

malware detection was 94.0%, and the accuracy for two major malware families was 81.8%. In addition, CNN-based methods can improve efficiency by reduce the size of the network. This further optimization can increase the applicability of the solution to IoT with fewer computing resources.

5) OPCODE AND GRAPH

Opcode can be a suitable and reliable feature for identifying malware using ML. With the attempts and experiments of many researchers, the combination of Windows malware opcodes and ML can effectively detect malware. Azmoodeh *et al.* [83] converted the selected features (opcodes) of each sample (software) into a graph (see Fig. 10). In their graph, nodes represented the opcodes and edges represented the affinity of each node (which needs to be calculated) in the disassembly file of each software. Graphs can be converted into eigenspace [84], so that CNN can be used to classify the generated malicious and benign software graphs. In Azmoodeh's experiment, the opcode sequences of 1078 benign software and 128 malwares were extracted. Using graphs converted from opcodes, the detection accuracy of malware samples was 99.68%, and the recall rate was 98.37%. This method is very effective for malware identification.

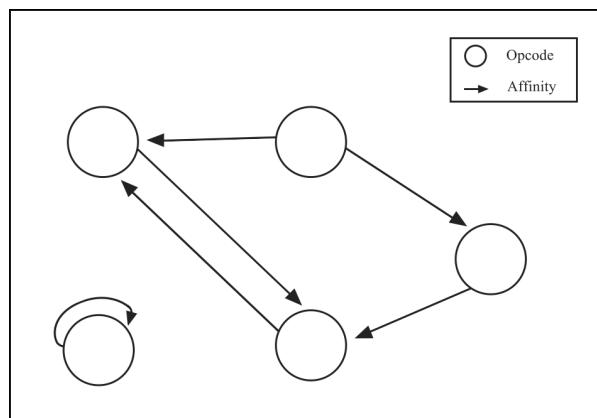


FIGURE 10. A graph converted from opcodes.

6) ACTIVE IMMUNITY FOR MALWARE DETECTION

In addition to extracting different forms of software information and using algorithms to detect malware, the recent development of adversarial machine learning provides a new way to enhance the active immunity of malware detection. When we use ML algorithms, we should always pay attention to the fact that ML may also be cheated by attackers. The attacker will avoid the detection according to the weakness of ML. If the detector can predict the possible evasion choice of the attacker in advance, it can greatly reduce the detection delay and loss caused by unknown attacks and form “active immunity”.

Chen *et al.* [85] described this as “arms race between evasion attack and defense”. First of all, they proposed an

effective evasion model (EvnAttack) by simulating attackers. Then, in order to effectively combat this kind of evasion attack, they further proposed a malware detection learning paradigm (SecDefender), which considered the cost of attacker's evasion attack. The effectiveness of this method was proved by comprehensive experiments on the real data sets of Comodo cloud security center. Wu *et al.* [86] pointed out that malware can avoid ML detection by constantly modifying its structure while maintaining malicious behaviors. Reinforcement learning can continuously simulate attackers to generate new malware samples, thus providing possible attack references for defenders. They designed a model based on reinforcement learning to improve the possibility of these newly generated malware to evade ML model, and then retrained the detection model by using these newly generated samples. In their experiments, the detection accuracy of malware was improved from 15.75% to 93.5% after retraining, which greatly improved the ability of the detection model for unknown attacks. Demontis *et al.* [87] studied the existing attack frameworks, summarized and classified current attacker's targets, knowledge, attack modes and potential attack scenarios. Then they implemented a set of evasion attacks to evaluate the security of malware detectors. They pointed out that linear and nonlinear classifiers with uniformly distributed feature weights can improve the system security without significantly affecting the computational efficiency.

V. CHALLENGES AND FUTURE DIRECTIONS

Artificial intelligence has put forward many effective solutions to solve the security problems of IoT, but this does not mean that IoT security has been properly solved. There are still many challenges to be faced in data, algorithm, and architecture.

A. DATA CHALLENGES

Data is the basis for the application of AI methods in the field of IoT. The heterogeneity of IoT makes a large number of data generated in different fields, which may lead to various problems such as poor data availability and quality, hidden dangers of data privacy, difficulties in data integration and so on.

1) THE AVAILABILITY OF DATA SETS

Machine learning and deep learning require a large number of training data sets. Most deep learning methods are usually based on high-quality data to achieve their good performance [88]. Comprehensive and diverse training data is the basis for the model to acquire knowledge. The quality of training data sets will directly affect the effectiveness of the model, so it is necessary to include as much as possible attack information that reflects the real world. In addition to training data, the model also requires test data sets to analyze and compare the generalization capabilities of various algorithms, so as to evaluate and improve training models. Therefore, in the context of applying machine learning and deep learning

to IoT security, how to obtain high availability training and test data sets containing various possible attack types is an important challenge.

2) MAINTAIN THE QUALITY OF NEW DATA IN REAL TIME

Due to the rapid data generation speed and the complexity of data sources of IoT, maintaining high-quality data in real time is a challenging task [18]. The newly acquired data should be of the same quality as the original data set. Otherwise, with the addition of new data, the feature space of the original training model will be gradually destroyed, resulting in model failure. Although some solutions have been proposed to solve these problems, due to the decentralization of big data management, no solution can handle all aspects of data.

3) DATA CLEANING [89]

Data cleaning is the process of removing duplicate information, correcting existing errors, and providing data consistency. It is estimated that anomaly and impurity in the data generally account for about 5% of the total data, which may be even worse for IoT. The data types that need to be cleaned are:

- Incomplete Data. Incomplete data is data with missing information, such as missing fields. The incomplete data needs to be filtered out, supplemented or abandoned according to the actual situation. Only the complete data can be written into database. In most cases, missing values need to be filled in manually. Some missing values can also be derived from data source and can be replaced by average, maximum, minimum, or more complex probability estimates.
- Incorrect Data. The reason for incorrect data is that the business system is not sound enough. The data is written directly into database without judging whether the data is correct when input, such as the value is out of bounds and the date format is incorrect. As with incomplete data, incorrect data needs to be corrected and reviewed before it can be written into database again. Statistical analysis can identify possible wrong or abnormal values, for example, deviation analysis can identify values that do not conform to the distribution or regression equation. Simple rules (common-sense rules, business specific rules, etc.) or constraints between attributes can also detect incorrect data.
- Duplicate Data. Records with the same attribute value in the database are considered as duplicate data. It is necessary to merge same records into one record to avoid affecting the use efficiency.

4) DATA INTEGRATION [90]

IoT obtains massive data from different types of sensor devices such as RFID, ZigBee, GPS, which need to be properly integrated before they can be used. Different data sources often have heterogeneous data, so how to integrate heterogeneous data while ensuring data quality is a challenge. Data integration is to integrate data of different sources,

formats, and properties logically or physically to provide comprehensive data sharing for users. Different types of data generated in various fields of IoT can be divided into three categories: (a) structured data, such as table data with rows and columns stored in traditional database systems; (b) semi-structured data, such as HTML, XML files; (c) unstructured data, such as videos and images. There are already some frameworks available in the field of enterprise data integration. Currently, federal-based, middleware-based and data warehouse methods are used to construct integrated systems. These technologies solve data integration and data sharing in different focuses and applications, providing decision support for enterprises. However, the applicability of these technologies for IoT data need to be investigated, and appropriate improvements are required.

5) DATA SECURITY AND PRIVACY

Data security and privacy are challenges in processing and using data. In the process of data collection and transmission, you may face the risk of leakage of privacy. Although data encryption provides enhanced privacy protection, many users may question their security because the system does not provide a reliable service level agreement (SLA) regarding the theft or misuse of personal information of users [91]. For example, personal information contained in wearable device data can easily endanger users' privacy. Personalized medical and healthcare applications rely on human body data collected from wearable devices for medical diagnosis and service recommendations. These personally related information is very rich, usually including location, identity and physiological characteristics. It's easy to infer individual habits, behaviors, and preferences. Personal information in some areas such as medical care must be carefully protected by all parties involved in data acquisition, management, and utilization [92].

B. ALGORITHM CHALLENGES

Artificial intelligence is not omnipotent, but also has many defects. When using ML to solve the security problems of IoT, the defects are inherited into IoT, which needs attention and improvement.

1) WEAK PORTABILITY

ML models are always specific to a certain field. When a good model is obtained for a scene and transferred to other similar problems, the original model parameters may fail. We need to retrain new models to replace the original one. For IoT, the diversification of applications will correspond to many different models, all of which need to be maintained and updated. But the process of retraining is cumbersome, there will be many unknown errors, resulting in a huge waste of time and computing resources.

2) INSTABILITY [93]

In ML and DL, small changes in the input may cause different effects on the output. Even if the input data of models changes

slightly, the output may change dramatically. Attackers can deliberately change some input data, resulting in poor system stability and unexpected results. Therefore, it is important to maintain the integrity and stability of the input data, but it is not an easy task in IoT environment that generates a large amount of high-frequency data.

3) POOR INTERPRETABILITY / LOW TRANSPARENCY / BLACK BOX

Deep neural network models are like black boxes because we have no way of knowing how they draw conclusions by manipulating parameters and input data [94]. There are two opposite trends at the same time: the increasing number of network layers and the decreasing interpretability. When we want to improve the effect of the model, it is inevitable to increase the number of network layers, resulting in a sharp increase in the complexity of the model and poor interpretability. Poor interpretability and low transparency make it difficult to find errors when the model fails, which is a serious problem for some high-risk areas. For example, when CNN is used in disease diagnosis of medical image, it is necessary to make reliable reasoning for its prediction to ensure the accuracy of diagnosis, but the lack of interpretability makes it difficult to apply the model to these work scenarios, reducing the actual benefits of the model.

4) COMPUTATIONAL COMPLEXITY AND RESOURCE CONSUMPTION

The computational complexity [95] and resource consumption of ML and DL are in sharp contrast to IoT devices. IoT devices are resource-constrained devices, and the acquisition of memory and computing resources is extremely limited. However, even if a lot of computing resources are given to some models related to image, speech and natural language processing, it will still take days or even weeks to complete the training. Therefore, the development of ML and DL frameworks which can effectively reduce the computational complexity is considerable to provide effective security mechanisms. Especially for large-scale IoT systems, reducing computational complexity and resource consumption has important practical significance in future research.

5) ML / DL SECURITY RISKS

The AI methods used to maintain the security of the IoT also have different degrees of security risks, just like IoT itself. The potential threats of the attacker against AI include poisoning, evasion, impersonation, and reverse attacks [96]. Poisoning, evasion, and impersonation attacks change the training data by generating wrong label samples, maliciously modifying samples, simulating samples, so that the model learns wrong and invalid knowledge from the training data, reducing the classifier's ability to distinguish normal and abnormal behaviors, leading to the failure of the detection function of models. Reverse attacks use the application program interface (API) provided by the existing ML systems to collect some basic information about the target model, and

reverse analysis of the basic information to use the target model to obtain private data, such as patient medical data. When these attack methods are combined with multiple IoT services, there will be serious consequences.

C. ARCHITECTURAL CHALLENGES

When AI technologies are applied to IoT, they must face new trends in the development of IoT: mobile and distributed.

The current IoT network and services rely heavily on the “Cloud-Channel-Device” architecture. The establishment of large-scale cloud computing center can store and process a large amount of data in a centralized way, so as to use the computing power of massive machines in the data center to calculate and make decisions. Then, the analysis results are returned to the device to achieve the interconnection effect. The transmission, storage and processing of data also depend heavily on the C/S service architecture to process and respond all requests and instructions through the central server.

The cloud service realizes the construction of super computing and storage capabilities, which solves problems of high cost of infrastructure construction and low utilization rate of computing storage resources for small and medium-sized enterprises. However, in the context of IoT, this is inappropriate. The number of device connections and data generation increase exponentially, which brings the following challenges to the cloud architecture:

- The linear growth of centralized cloud computing capabilities cannot match the exponential growth of data generated by terminals.
- Mass data transmission to the cloud computing center dramatically increases the load of the transmission bandwidth, resulting in a large network delay, which poses severe challenges for delay-sensitive application scenarios (such as driverless cars, industrial manufacturing, etc.).
- A large amount of power consumption caused by data transmission brings a great burden to the cloud service.

The requirements for network transmission, data storage and high-performance computing force us to carry out data cleaning, processing and decision-making at the source of the data. The new architecture based on edge will become an important direction for the future development of IoT architecture. Distributed and edge architecture will completely change the original data application, which will put more stringent requirements on the application of AI in the field of IoT security.

D. FUTURE DIRECTIONS

Some of the above-mentioned challenges are inherent in IoT, such as poor data availability and the need for data integration. Some are the new challenges that AI brings to IoT, such as algorithm security and resource consumption. For these hidden dangers, we need to improve the existing technologies or develop new technologies. In addition to addressing these challenges, we point out two possible new directions here.

1) EMBEDDING EDGE AI CHIPS INTO IoT DEVICES

Most of traditional AI computing tasks are performed remotely on centralized core devices or platforms, but this is not the best solution for IoT. Edge AI chips make it possible to embed AI computations into IoT devices. Edge AI chip [97] is a chip that can perform or accelerate machine learning tasks on edge devices. At present, Google, NVIDIA, Intel, Qualcomm and Huawei are all rapidly developing edge AI chip technologies, such as Google Coral Edge TPU [98].

In many industrial fields, network conditions are not good and the cost of upgrading communication infrastructure is very high. Edge AI chips enable terminals to perform AI computations locally, which greatly reduces transmission costs. Edge AI chips can also greatly reduce delays. Edge computing has real-time requirements, because it is necessary to make real-time decisions on various devices. However, cloud computing and data center computing will have network delay, so it is difficult to achieve real-time performance.

Edge AI chips can also protect data privacy and security. The calculations don't need to send original data back to the cloud, which can greatly protect the security and privacy of data, reduce the possibility of data leakage and interception or abuse of personal / corporate data.

2) DEVELOPING SERVICE-ORIENTED IoT SECURITY ARCHITECTURES WHICH CAN MEET THE NEEDS OF DIFFERENT SERVICES IN DIFFERENT FIELDS

IoT security needs to meet the differentiated security needs of different application scenarios. How to design flexible security architectures to provide adaptive and differentiated security guarantee capabilities for industrial applications is an urgent problem. The key is that the network infrastructure can support the open ability of Security-as-a-Service [99], [100].

Network architectures need to establish security resources independent of devices and applications based on computing resources, such as authentication protocol, cryptographic algorithm, data encryption and decryption, etc. On the basis of security resources, we can use these resources to establish security functions such as trusted authentication, digital identity, operation maintenance and management. Then, we should build platforms using these security functions to provide security services to third-party applications through open APIs. At last, in the face of different IoT industries with different security requirements, the third party can flexibly use the security capabilities and services provided by the open platform to realize customized security protection. Compared with the third-party self-designed security solutions, such architecture design can achieve complete security protection, and can embed feasible AI models in open platforms, which greatly enhances the security capability of the third party. At the same time, such design also has strong scalability, so the security of IoT can also obtain a strong elastic security capability.

VI. CONCLUSION

The research of this article proves that AI is feasible for the security of IoT, especially for the four key risks: device authentication, DoS / DDoS attack defense, intrusion detection and malware detection. The general process of the AI schemes proposed by us can also be used as a reference to solve IoT security problems in the future. In addition, when AI is applied for IoT security, potential challenges in data, algorithm and architecture need to be solved to avoid adding new threats to IoT security. How to solve these challenges can serve as potential future research directions.

ACKNOWLEDGMENT

(*Hui Wu and Haiting Han contribute equally to this work.*)

REFERENCES

- [1] *ITU Internet Reports 2005: The Internet of Things*, Geneva, Switzerland: International Telecommunication Union, 2005.
- [2] C. Hai-ming, "Key Technologies and Applications of Internet of Things," *Comput. Sci.*, vol. 36, no. 6, pp. 1–4, 2010.
- [3] C.-L. Zhong, Z. Zhu, and R.-G. Huang, "Study on the IOT architecture and gateway technology," in *Proc. 14th Int. Symp. Distrib. Comput. Appl. Bus. Eng. Sci. (DCABES)*, Aug. 2015, pp. 196–199, doi: [10.1109/DCABES.2015.56](https://doi.org/10.1109/DCABES.2015.56).
- [4] M. Bauer, M. Boussard, N. Bui, J. D. Loof, C. Magerkurth, S. Meissner, A. Netsträter, J. Stefa, M. Thoma, and J. W. Walewski, "IoT reference architecture," in *Enabling Things to Talk*, 2013, pp. 163–211, doi: [10.1007/978-3-642-40403-0_8](https://doi.org/10.1007/978-3-642-40403-0_8).
- [5] J. Deogirikar and A. Vidhate, "Security attacks in IoT: A survey," in *Proc. Int. Conf.*, Feb. 2017, pp. 32–37.
- [6] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016, doi: [10.1109/JIOT.2016.2579198](https://doi.org/10.1109/JIOT.2016.2579198).
- [7] E. Bertino, "Data security and privacy in the IoT," in *Proc. EDBT*, 2016, pp. 1–3.
- [8] A. Singla, A. Mudgerkar, I. Papapanagiotou, and A. A. Yavuz, "HAA: hardware-accelerated authentication for Internet of Things in mission critical vehicular networks," in *Proc. MILCOM - IEEE Mil. Commun. Conf.*, Oct. 2015, pp. 1298–1304, doi: [10.1109/MILCOM.2015.7357624](https://doi.org/10.1109/MILCOM.2015.7357624).
- [9] *Cyber Security for Consumer Internet of Things*, document TS 103 645, ETSI, 2019.
- [10] W. U. Chuankun, L. Zhang, and L. I. Jiangli, "Design of trust architecture and lightweight authentication scheme for IoT devices," *Netinfo Secur.*, vol. 17, no. 9, pp. 16–20, Oct. 2017.
- [11] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, Aug. 2008, doi: [10.1016/j.comnet.2008.04.002](https://doi.org/10.1016/j.comnet.2008.04.002).
- [12] D. Coss, "The cia strikes back: Redefining confidentiality, integrity and availability in security," *J. Inf. Syst. Secur.*, vol. 10, no. 3, pp. 21–45, Jul. 2014. [Online]. Available: <http://www.proso.com/dl/Samonas.pdf>
- [13] C. Koliaris, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017, doi: [10.1109/MC.2017.201](https://doi.org/10.1109/MC.2017.201).
- [14] S. Al-Sarawi, M. Anbar, K. Alieyan, and M. Alzubaidi, "Internet of Things (IoT) communication protocols: Review," in *Proc. 8th Int. Conf. Inf. Technol. (ICIT)*, May 2017, pp. 685–690.
- [15] R. S. Sinha, Y. Wei, and S.-H. Hwang, "A survey on LPWA technology: LoRa and NB-IoT," *ICT Exp.*, vol. 3, no. 1, pp. 14–21, Mar. 2017, doi: [10.1016/j.icte.2017.03.004](https://doi.org/10.1016/j.icte.2017.03.004).
- [16] D. Gislason, *Zigbee Wireless Networking*. Oxford, U.K.: Newnes, 2008, pp. 3–14.
- [17] N. Chou, R. Ledesma, Y. Teraguchi, and J. C. Mitchell, "Client-side defense against Web-based identity theft," in *Proc. Netw. Distrib. Syst. Secur. Symp.*, Feb. 2004, p. 15. [Online]. Available: http://simson.net/ref/2005/csci_e-170/ref/webspoof.pdf
- [18] M. A. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Commun. Surveys Tuts.*, early access, Apr. 20, 2020, doi: [10.1109/COMST.2020.2988293](https://doi.org/10.1109/COMST.2020.2988293).

- [19] R. Dorai and V. Kannan, "SQL injection-database attack revolution and prevention," *J. Int. Commer. Law Technol.*, vol. 6, no. 4, pp. 224–231, 2011, doi: [10.4028/www.scientific.net/AMM.740.810](https://doi.org/10.4028/www.scientific.net/AMM.740.810).
- [20] Q. Jing, A. V. Vasilakos, J. Wan, J. Lu, and D. Qiu, "Security of the Internet of Things: Perspectives and challenges," *Wireless Netw.*, vol. 20, no. 8, pp. 2481–2501, Nov. 2014, doi: [10.1007/S11276-014-0761-7](https://doi.org/10.1007/S11276-014-0761-7).
- [21] P. Bisht and V. N. Venkatakrishnan, "XSS-GUARD: Precise dynamic prevention of cross-site scripting attacks," in *Proc. Int. Conf. Detection*, 2008, pp. 23–43, doi: [10.1007/978-3-540-70542-0_2](https://doi.org/10.1007/978-3-540-70542-0_2).
- [22] P. Faruki, A. Bharmal, V. Laxmi, V. Ganmoor, M. S. Gaur, M. Conti, and M. Rajarajan, "Android security: A survey of issues, malware penetration, and defenses," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 998–1022, 2nd Quart., 2015, doi: [10.1109/COMST.2014.2386139](https://doi.org/10.1109/COMST.2014.2386139).
- [23] R. Doshi, N. Aphorpe, and N. Feamster, "Machine learning DDoS detection for consumer Internet of Things devices," in *Proc. IEEE Secur. Privacy Workshops (SPW)*, May 2018, pp. 29–35, doi: [10.1109/SPW.2018.00013](https://doi.org/10.1109/SPW.2018.00013).
- [24] A. Madaan, X. Wang, W. Hall, and T. Tiropanis, "Observing data in IoT worlds: What and how to observe?" in *Proc. Living Internet Things, Cybersecurity (IoT)*, London, U.K., 2018, pp. 1–7, doi: [10.1049/cp.2018.0032](https://doi.org/10.1049/cp.2018.0032).
- [25] A. Riahi, Y. Challal, E. Natalizio, Z. Chtourou, and A. Bouabdallah, "A systemic approach for IoT security," in *Proc. IEEE Int. Conf. Distrib. Comput. Sensor Syst.*, May 2013, pp. 351–355.
- [26] M. S. Mahdavinejad, M. Rezvan, M. Barekatian, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for Internet of Things data analysis: A survey," *Digit. Commun. Netw.*, vol. 4, no. 3, pp. 161–175, Aug. 2018, doi: [10.1016/J.DCAN.2017.10.002](https://doi.org/10.1016/J.DCAN.2017.10.002).
- [27] R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," in *Proc. ACM Int. Conf. Proc. Ser.*, 2006, pp. 161–168, doi: [10.1145/1143844.1143865](https://doi.org/10.1145/1143844.1143865).
- [28] B. C. Love, "Comparing supervised and unsupervised category learning," *Psychonomic Bull. Rev.*, vol. 9, no. 4, pp. 829–835, Dec. 2002, doi: [10.3758/BF03196342](https://doi.org/10.3758/BF03196342).
- [29] K. Gai and M. Qiu, "Optimal resource allocation using reinforcement learning for IoT content-centric services," *Appl. Soft Comput.*, vol. 70, pp. 12–21, Sep. 2018, doi: [10.1016/j.asoc.2018.03.056](https://doi.org/10.1016/j.asoc.2018.03.056).
- [30] R. Polikar, "Ensemble Learning," *Ensemble Mach. Learn.*, vol. 4, pp. 1–34, Jan. 2012, doi: [10.1007/978-1-4419-9326-7_1](https://doi.org/10.1007/978-1-4419-9326-7_1).
- [31] M. Mohammadi, A. Al-Fuqaha, S. Sorour, and M. Guizani, "Deep learning for IoT big data and streaming analytics: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2923–2960, 1st Quart., 2018, doi: [10.1109/COMST.2018.2844341](https://doi.org/10.1109/COMST.2018.2844341).
- [32] G. Liang, "Automatic traffic accident detection based on the Internet of Things and support vector machine," *Int. J. Smart Home*, vol. 9, no. 4, pp. 97–106, Apr. 2015, doi: [10.14257/ijsh.2015.9.4.10](https://doi.org/10.14257/ijsh.2015.9.4.10).
- [33] L. Bilge and T. Dumitra, "Before we knew it: An empirical study of zero-day attacks in the real world," in *Proc. Comput. Commun. Secur.*, 2012, pp. 833–844, doi: [10.1145/2382196.2382284](https://doi.org/10.1145/2382196.2382284).
- [34] A. K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, Jun. 2010, doi: [10.1016/j.patrec.2009.09.011](https://doi.org/10.1016/j.patrec.2009.09.011).
- [35] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [36] M. Bogdanoski, T. Shuminoski, and A. Risteski, "Analysis of the SYN flood DoS attack," *Int. J. Comput. Netw. Inf. Secur.*, vol. 5, pp. 1–11, Jun. 2013.
- [37] H. Fujinoki, "Cached Guaranteed-Timer Random-Drop against TCP SYN-flood Attacks and Flash Crowds," in *Proc. IASTED Int. Conf. Commun., Netw., Inf. Secur.*, vol. 2005, pp. 162–169.
- [38] A. N. Bhagoji, D. Cullina, C. Sitawarin, and P. Mittal, "Enhancing robustness of machine learning systems via data transformations," in *Proc. 52nd Annu. Conf. Inf. Sci. Syst. (CISS)*, Mar. 2018, pp. 1–5.
- [39] B. Neyshabur, S. Bhojanapalli, D. McAllester, and N. Srebro, "Exploring generalization in deep learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5947–5956.
- [40] R. Moore and J. DeNero, "L1 and L2 regularization for multiclass hinge loss models," in *Proc. Symp. Mach. Learn. Speech Lang. Process.*, 2011, p. 15.
- [41] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 660–674, 1991, doi: [10.1109/21.97458](https://doi.org/10.1109/21.97458).
- [42] K. P. Murphy, *Naive Bayes Classifiers*. Vancouver, BC, Canada: University of British Columbia, 2006.
- [43] A. Elisseeff and J. Weston, "A kernel method for multi-labelled classification," in *Proc. Adv. Neural Inf. Process. Syst.*, 2002, pp. 681–687.
- [44] R. Chettri, S. Pradhan, and L. Chettri, "Internet of Things: Comparative study on classification algorithms (k-NN, naive Bayes and case based Reasoning)," *Int. J. Comput. Appl.*, vol. 130, no. 12, pp. 7–9, Nov. 2015, doi: [10.5120/IJCA2015907120](https://doi.org/10.5120/IJCA2015907120).
- [45] M. Ringnér, "What is principal component analysis?" *Nature Biotechnol.*, vol. 26, no. 3, pp. 303–304, Mar. 2008, doi: [10.1038/nbt0308-303](https://doi.org/10.1038/nbt0308-303).
- [46] A. Liaw and M. Wiener. (2007). *Classification and Regression by Random Forest*. [Online]. Available: <http://cogns.northwestern.edu/cbmg/LiawAndWiener2002.pdf>
- [47] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, nos. 3–4, pp. 279–292, May 1992, doi: [10.1007/bf00992698](https://doi.org/10.1007/bf00992698).
- [48] S.-C. B. Lo, H.-P. Chan, J.-S. Lin, H. Li, M. T. Freedman, and S. K. Mun, "Artificial convolution neural network for medical image pattern recognition," *Neural Netw.*, vol. 8, nos. 7–8, pp. 1201–1214, 1995, doi: [10.1016/0893-6080\(95\)00061-5](https://doi.org/10.1016/0893-6080(95)00061-5).
- [49] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, 1997, doi: [10.1109/78.650093](https://doi.org/10.1109/78.650093).
- [50] Y. T. Zhang, C. H. Yan, and Y. R. Wei, "Research on security of IoT perception layer based on node authentication," *Netinf. Secur.*, vol. 15, no. 11, pp. 27–32, Apr. 2015.
- [51] R. K. C. Chang, "Defending against flooding-based distributed denial-of-service attacks: A tutorial," *IEEE Commun. Mag.*, vol. 40, no. 10, pp. 42–51, Oct. 2002, doi: [10.1109/MCOM.2002.1039856](https://doi.org/10.1109/MCOM.2002.1039856).
- [52] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 2046–2069, 2nd Quart., 2013, doi: [10.1109/SURV.2013.031413.00127](https://doi.org/10.1109/SURV.2013.031413.00127).
- [53] B. Mukherjee, L. T. Heberlein, and K. N. Levitt, "Network intrusion detection," *IEEE Netw.*, vol. 8, no. 3, pp. 26–41, May 1994, doi: [10.1109/65.283931](https://doi.org/10.1109/65.283931).
- [54] A. Moser, C. Kruegel, and E. Kirda, "Exploring multiple execution paths for malware analysis," in *Proc. IEEE Symp. Secur. Privacy*, Oct. 2007, pp. 231–245, doi: [10.1109/SP.2007.17](https://doi.org/10.1109/SP.2007.17).
- [55] M. R. Hasan, M. Jamil, and M. Rahman, "Speaker identification using mel frequency cepstral coefficients," *Variation*, vol. 1, no. 4, p. 25, 2004.
- [56] H. Jabbar and R. Z. Khan, "Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study)," in *Proc. Comput. Sci., Commun. Instrum. Devices*, 2015, pp. 163–172.
- [57] C. Rijsbergen, *Information Retrieval*. Newton, MA, USA: Butterworth-Heinemann, 1979.
- [58] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognit.*, vol. 30, no. 7, pp. 1145–1159, 1997, doi: [10.1016/S0031-3203\(96\)00142-2](https://doi.org/10.1016/S0031-3203(96)00142-2).
- [59] C. Adams, P. Sylvester, M. Zolotarev, and R. Zuccherato, "Internet X. 509 public key infrastructure data validation and certification server protocols," *Request Comments*, vol. 3029, p. 15, Oct. 2001.
- [60] A. Shimizu, T. Horioka, and H. Inagaki, "A Password Authentication Method for Contents Communications on the Internet," *IEICE Trans. Commun.*, vol. 81, no. 8, pp. 1666–1673, 1998.
- [61] Z. Li, W. Zuoyue, W. Chundong, M. A. Yunfei, and X. Chaocan, "Design and implementation of intelligent identification system for IoT terminals," *J. Chongqing Univ. Posts Telecommun.*, vol. 31, no. 4, pp. 443–450, 2019.
- [62] Y. Meidan, M. Bohadana, A. Shabtai, M. Ochoa, N. O. Tippenhauer, J. D. Guarnizo, and Y. Elovici, "Detection of unauthorized IoT devices using machine learning techniques," 2017, *arXiv:1709.04647*. [Online]. Available: <http://arxiv.org/abs/1709.04647>
- [63] J. Chauhan, S. Seneviratne, Y. Hu, A. Misra, A. Seneviratne, and Y. Lee, "Breathing-based authentication on resource-constrained IoT devices using recurrent neural networks," *Computer*, vol. 51, no. 5, pp. 60–67, May 2018, doi: [10.1109/MC.2018.2381119](https://doi.org/10.1109/MC.2018.2381119).
- [64] P. Punithavathi, S. Geetha, M. Karuppiah, S. H. Islam, M. M. Hassan, and K.-K.-R. Choo, "A lightweight machine learning-based authentication framework for smart IoT devices," *Inf. Sci.*, vol. 484, pp. 255–268, May 2019.
- [65] C. Shi, J. Liu, H. Liu, and Y. Chen, "Smart User authentication through actuation of daily activities leveraging WIFI-enabled IoT," in *Proc. ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2017, pp. 1–7, doi: [10.1145/3084041.3084061](https://doi.org/10.1145/3084041.3084061).
- [66] Y. Lin, X. Zhu, Z. Zheng, Z. Dou, and R. Zhou, "The individual identification method of wireless device based on dimensionality reduction and machine learning," *J. Supercomput.*, vol. 75, no. 6, pp. 3010–3027, Jun. 2019.

- [67] B. Bezwada, M. Bachani, J. Peterson, H. Shirazi, I. Ray, and I. Ray, "Behavioral fingerprinting of IoT devices," in *Proc. Workshop Attacks Solutions Hardw. Secur.*, 2018, pp. 41–50.
- [68] J. Ye, X. Cheng, J. Zhu, L. Feng, and L. Song, "A DDoS attack detection method based on SVM in software defined network," *Secur. Commun. Netw.*, vol. 2018, pp. 1–8, Oct. 2018, doi: [10.1155/2018/9804061](https://doi.org/10.1155/2018/9804061).
- [69] R. V. Kulkarni and G. K. Venayagamoorthy, "Neural network based secure media access control protocol for wireless sensor networks," in *Proc. Int. Joint Conf. Neural Netw.*, 2009, pp. 3437–3444, doi: [10.1109/IJCNN.2009.5179075](https://doi.org/10.1109/IJCNN.2009.5179075).
- [70] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Gener. Comput. Syst.*, vol. 82, pp. 761–768, May 2018.
- [71] C. Y. Chung, M. Gertz, and K. Levitt, "DEMIDS: A misuse detection system for database systems," in *Proc. Working Conf. Integrity Internal Control Inf. Syst.*, 2000, pp. 159–178, doi: [10.1007/978-0-387-35501-6_12](https://doi.org/10.1007/978-0-387-35501-6_12).
- [72] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surveys*, vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: [10.1145/1541880.1541882](https://doi.org/10.1145/1541880.1541882).
- [73] G. Kim, S. Lee, and S. Kim, "A novel hybrid intrusion detection method integrating anomaly detection with misuse detection," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1690–1700, Mar. 2014, doi: [10.1016/J.ESWA.2013.08.066](https://doi.org/10.1016/J.ESWA.2013.08.066).
- [74] J. Li, Z. Zhao, R. Li, and H. Zhang, "AI-based two-stage intrusion detection for software defined IoT networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2093–2102, Apr. 2019.
- [75] M.-Y. Su, "Real-time anomaly detection systems for Denial-of-Service attacks by weighted k-nearest-neighbor classifiers," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3492–3498, Apr. 2011, doi: [10.1016/J.ESWA.2010.08.137](https://doi.org/10.1016/J.ESWA.2010.08.137).
- [76] H. H. W. J. Bosman, G. Iacca, A. Tejada, H. J. Wörtche, and A. Liotta, "Ensembles of incremental learners to detect anomalies in ad hoc sensor networks," *Ad Hoc Netw.*, vol. 35, pp. 14–36, Dec. 2015, doi: [10.1016/J.ADHOC.2015.07.013](https://doi.org/10.1016/J.ADHOC.2015.07.013).
- [77] D. Venugopal and G. Hu, "Efficient signature based malware detection on mobile devices," *Mobile Inf. Syst.*, vol. 4, no. 1, pp. 33–49, 2008.
- [78] J. Scott, "Signature based malware detection is dead," Inst. Crit. Infrastruct. Technol., Washington, DC, USA, White Paper, 2017. [Online]. Available: <https://icitech.org/wp-content/uploads/2017/02/ICIT-Analysis-Signature-Based-Malware-Detection-is-Dead.pdf>
- [79] W. Zhou and B. Yu, "A cloud-assisted malware detection and suppression framework for wireless multimedia system in IoT based on dynamic differential game," *China Commun.*, vol. 15, no. 2, pp. 209–223, Feb. 2018, doi: [10.1109/CC.2018.8300282](https://doi.org/10.1109/CC.2018.8300282).
- [80] S. Y. Yerima, I. Muttil, and S. Sezer, "High accuracy Android malware detection using ensemble learning," *IET Inf. Secur.*, vol. 9, no. 6, pp. 313–320, Nov. 2015, doi: [10.1049/IET-IFS.2014.0099](https://doi.org/10.1049/IET-IFS.2014.0099).
- [81] D. Zhu, H. Jin, Y. Yang, D. Wu, and W. Chen, "DeepFlow: Deep learning-based malware detection by mining Android application for abnormal usage of sensitive data," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Jul. 2017, pp. 438–443.
- [82] J. Su, V. D. Vasconcellos, S. Prasad, S. Daniele, Y. Feng, and K. Sakurai, "Lightweight classification of IoT malware based on image recognition," *Comput. Softw. Appl. Conf.*, vol. 2, pp. 664–669, Oct. 2018, doi: [10.1109/COMPSSAC.2018.10315](https://doi.org/10.1109/COMPSSAC.2018.10315).
- [83] A. Azmoodeh, A. Dehghanianha, and K.-K.-R. Choo, "Robust malware detection for Internet of (Battlefield) Things devices using deep eigenspace learning," *IEEE Trans. Sustain. Comput.*, vol. 4, no. 1, pp. 88–95, Jan. 2019, doi: [10.1109/TSUSC.2018.280965](https://doi.org/10.1109/TSUSC.2018.280965).
- [84] F. R. Chung, *Spectral Graph Theory*, no. 92. American Mathematical Society, 1997.
- [85] L. Chen, Y. Ye, and T. Bourlai, "Adversarial machine learning in malware detection: Arms race between evasion attack and defense," in *Proc. Eur. Intell. Secur. Inform. Conf.*, 2017, pp. 99–106.
- [86] C. Wu, J. Shi, Y. Yang, and W. Li, "Enhancing machine learning based malware detection model by reinforcement learning," in *Proc. 8th Int. Conf. Commun. Netw. Secur.*, 2018, pp. 74–78.
- [87] A. Demontis, M. Melis, B. Biggio, D. Maiorca, D. Arp, K. Rieck, I. Corona, G. Giacinto, and F. Roli, "Yes, machine learning can be more secure! A case study on Android malware detection," *IEEE Trans. Dependable Secure Comput.*, vol. 16, no. 4, pp. 711–724, Jul. 2019.
- [88] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning for big data," *Inf. Fusion*, vol. 42, pp. 146–157, Jul. 2018, doi: [10.1016/J.INFFUS.2017.10.006](https://doi.org/10.1016/J.INFFUS.2017.10.006).
- [89] E. Rahm and H. H. Do, "Data cleaning: Problems and current approaches," *Eng. Bull.*, vol. 23, pp. 3–13, Oct. 2000. [Online]. Available: http://dc-pubs.dbs.uni-leipzig.de/files/Rahm2000Data/Cleaning_Problemsand.pdf
- [90] M. Lenzerini, "Data integration: A theoretical perspective," in *Proc. Symp. Princ. Database Syst.*, 2002, pp. 233–246, doi: [10.1145/543613.543644](https://doi.org/10.1145/543613.543644).
- [91] M. Marjani, F. Nasaruddin, A. Gani, A. Karim, I. Abaker Targio Hashem, A. Siddiq, and I. Yaqoob, "Big IoT data analytics: Architecture, opportunities, and open research challenges," *IEEE Access*, vol. 5, pp. 5247–5261, 2017, doi: [10.1109/ACCESS.2017.2689040](https://doi.org/10.1109/ACCESS.2017.2689040).
- [92] E. Bertino, "Security and privacy in the IoT," in *Proc. Int. Conf. Inf. Secur. Cryptol.*, 2017, pp. 3–10, doi: [10.1007/978-3-319-75160-3_1](https://doi.org/10.1007/978-3-319-75160-3_1).
- [93] A. Kaplan, D. J. Nordman, and S. B. Vardeman, "On the S-instability and degeneracy of discrete deep learning models," *Inf. Inference*, vol. 12, pp. 1–29, Nov. 2019, doi: [10.1093/imaiia/fiaz022](https://doi.org/10.1093/imaiia/fiaz022).
- [94] D. Castelvecchi, "Can we open the black box of AI?" *Nature*, vol. 538, no. 7623, pp. 20–23, Oct. 2016, doi: [10.1038/538020A](https://doi.org/10.1038/538020A).
- [95] H. Huang, Y. Song, J. Yang, G. Gui, and F. Adachi, "Deep-Learning-Based millimeter-wave massive MIMO for hybrid precoding," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 3027–3032, Mar. 2019, doi: [10.1109/TVT.2019.2893928](https://doi.org/10.1109/TVT.2019.2893928).
- [96] Q. Liu, P. Li, W. Zhao, W. Cai, S. Yu, and V. C. M. Leung, "A survey on security threats and defensive techniques of machine learning: A data driven view," *IEEE Access*, vol. 6, pp. 12103–12117, 2018, doi: [10.1109/ACCESS.2018.2805680](https://doi.org/10.1109/ACCESS.2018.2805680).
- [97] Z. Zhou, X. Chen, E. Li, L. Zeng, K. Luo, and J. Zhang, "Edge intelligence: Paving the last mile of artificial intelligence with edge computing," *Proc. IEEE*, vol. 107, no. 8, pp. 1738–1762, Aug. 2019.
- [98] J. Sengupta, R. Kubendran, E. Neftci, and A. Andreou, "High-speed, real-time, spike-based object tracking and path prediction on Google edge TPU," in *Proc. 2nd IEEE Int. Conf. Artif. Intell. Circuits Syst. (AICAS)*, Aug. 2020, pp. 134–135.
- [99] I. Hafeez, A. Y. Ding, L. Suomalainen, S. Häntönen, V. Niemi, and S. Tarkoma, "Cloud-based security as a service for smart IoT environments," in *Proc. 2015 Workshop Wireless Students*, 2015, p. 20.
- [100] T. Bhattachari and N. Chaki, "Poster: Exploring security as a service for IoT enabled remote application framework," in *Proc. 14th Annu. Int. Conf. Mobile Syst., Appl., Services Companion*, 2016, p. 15.



HUI WU received the bachelor's degree in software engineering from Beijing Jiaotong University, Beijing, China, in 2018. He is currently pursuing the master's degree with the School of Software and Microelectronics, Peking University. His research interests include explainable machine learning, safe and trusted AI, and the IoT.



HAITING HAN (Member, IEEE) received the M.S. degree in software engineering from Peking University, Beijing, in 2019. He was looking forward to find the connection between Protocol, Economy Institutions, Cyber Governance (Resource/Powers Distribution), and Economic Performance. From 2015 to 2016, he was an Analyst with Tsinghua Technology and Innovation Holdings Company, Ltd. From 2018 to 2020, he was a Researcher with the China Academy of Information and Communications Technology. He is currently a Ph.D. Fellow with the Center of Blockchain and Electronic Markets, University of Copenhagen. He is also attempting design a new economic system based on distributed technology and smart contracts—Data Market Infrastructures (DMIs) and E-Market Simulator System (EMSS). His research interests include cyber-physical systems (CPS) and distributed economics design.



XIAO WANG received the bachelor's degree in computer science and technology from Jilin University, Jilin, China, in 2018. He is currently pursuing the master's degree in big data and artificial intelligence with the School of Software and Microelectronics, Peking University. He is also an Intern with the Institute of Computing, Chinese Academy of Sciences. His research interests include machine learning, AI, the Internet of Things, and big data.



SHENGLI SUN received the Ph.D. degree from Fudan University, Shanghai, China, in 2008. He is currently an Associate Professor with the School of Software and Microelectronics, Peking University, Beijing, China. He has published more than 50 research articles in journals and conferences. His research interests include database, data mining, machine learning, and their applications in social networks analysis and intelligent diagnosis.

• • •