Contents lists available at ScienceDirect

Digital Communications and Networks

journal homepage: www.keaipublishing.com/dcan

Application of meta-learning in cyberspace security: a survey

Aimin Yang[a,b], Chaomeng Lu[a,c], Jie Li[b,*], Xiangdong Huang[a,c], Tianhao Ji[c], Xichang Li[c], Yichao Sheng[c]

[a] *Hebei Key Laboratory of Data Science and Application, North China University of Science and Technology, Tangshan, 063000, China*
[b] *Hebei Engineering Research Center for the Intelligentization of Iron Ore Optimization and Ironmaking Raw Materials Preparation Processes, North China University of Science and Technology, Tangshan, 063000, China*
[c] *The Key Laboratory of Engineering Computing in Tangshan City, North China University of Science and Technology, Tangshan, 063000, China*

ARTICLE INFO

Keywords:
Meta-learning
Cyberspace security
Machine learning
Few-shot learning

ABSTRACT

In recent years, machine learning has made great progress in intrusion detection, network protection, anomaly detection, and other issues in cyberspace. However, these traditional machine learning algorithms usually require a lot of data to learn and have a low recognition rate for unknown attacks. Among them, “one-shot learning”, “few-shot learning”, and “zero-shot learning” are challenges that cannot be ignored for traditional machine learning. The more intractable problem in cyberspace security is the changeable attack mode. When a new attack mode appears, there are few or even zero samples that can be learned. Meta-learning comes from imitating human problem-solving methods as humans can quickly learn unknown things based on their existing knowledge when learning. Its purpose is to quickly obtain a model with high accuracy and strong generalization through less data training. This article first divides the meta-learning model into five research directions based on different principles of use. They are model-based, metric-based, optimization-based, online-learning-based, or stacked ensemble-based. Then, the current problems in the field of cyberspace security are categorized into three branches: cyber security, information security, and artificial intelligence security according to different perspectives. Then, the application research results of various meta-learning models on these three branches are reviewed. At the same time, based on the characteristics of strong generalization, evolution, and scalability of meta-learning, we contrast and summarize its advantages in solving problems. Finally, the prospect of future deep application of meta-learning in the field of cyberspace security is summarized.

1. Introduction

There has been a frequent occurrence of various security incidents and cyber attacks in cyberspace. For example, the ransomware WannaCry in May 2017 used system vulnerabilities to attack, causing computer poisoning of hundreds of thousands of users in many countries. On the 9th of May 2021, the United States suddenly declared a state of emergency after the country's largest fuel pipeline operator was taken offline by a cyber-attack. Cyberspace not only includes hardware and software, such as the Internet, communication networks, various computing systems, various embedded processors and controllers, but also various data or information generated, processed, transmitted, and stored by these hardware and software, as well as the impact of human activities in it. Therefore, cyberspace is called the fifth-largest space besides land, sea, air, and space [1]. In the process of rapid development, the Internet gradually has the characteristics of large-scale, high-throughput, multiple connections, and so on, all leading to the complexity of the cyberspace environment. The rapid development of Internet services has also led to a substantial increase in cyber attacks. Cyber threats are becoming more and more complex, and the degree of automation is getting higher and higher, making many protective measures ineffective. Traditional cyberspace security methods have limited effectiveness in dealing with new cyber threats. Therefore, new methods are urgently needed.

In recent years, machine learning has achieved remarkable successes in dealing with various cyberspace security issues, including intrusion detection [2], malware detection and analysis, botnets, spam detection, etc. [3–5]. However, traditional machine learning models, such as Back Propagation Neural Network (BPNN), need to select the best activation function [6]. These models require multiple training to update iteratively, especially those that simulate high-complexity problems. When there are only a small number of training samples, even the timely

[*] Corresponding author.
E-mail addresses: aimin@ncst.edu.cn (A. Yang), lucm@stu.ncst.edu.cn (C. Lu), lijie@ncst.edu.cn (J. Li).

https://doi.org/10.1016/j.dcan.2022.03.007
Received 9 June 2021; Received in revised form 18 January 2022; Accepted 7 March 2022
Available online 17 March 2022

machine learning algorithm using gradient descent is not very effective [7]. In multi-classification using unbalanced data, the problem becomes more serious [8]. The super generalization ability of humans to face problems is still an unsolved problem for artificial intelligence, and meta-learning is considered as a strategy to overcome this challenge. Historically, the term meta-learning has been used in different fields. In the broadest sense, it encapsulates all systems that utilize previous learning experiences to learn new tasks faster [9]. This broad concept includes algorithm selection and hyperparameter optimization for traditional machine learning [10]. The key idea is that the main body of meta-learning improves its learning ability over time, namely, the ability of learning to learn. The learning process is mainly related to task sets and occurs at two different levels: internal and external. At the internal level, a new task is proposed to try to quickly learn the related concepts from training observations. This rapid adaptation benefits from the knowledge accumulated in early external tasks. Therefore, the inner layer focuses on a single task, while the outer layer focuses on multiple tasks. Based on the dual learning architecture of its inner and outer layers and the integrated collection of metadata, meta-learning can be used for zero-shot learning, one-shot learning, and few-shot learning, which can be used to address the current technical challenges in cyberspace security-related issues.

In the following sections, we first classify and introduce the application research of meta-learning in various fields of cyberspace security, then make a conclusive summary, and finally look forward to future research trends. In Section 2, meta-learning and its current research status in cyberspace security are introduced. Five research directions of meta-learning and three branches of cyberspace security are summarized, and the research framework of this article is given. Sections 3, 4, 5 respectively analyze and introduce the research progress of meta-learning applications in the sub-fields of cyberspace security. They are cyber security, information security, and artificial intelligence security. We also sort out the meta-learning methods and their innovations and analyze the advantages and disadvantages of various algorithms of meta-learning in different fields of cyberspace security. Then we give future prospects in combination with current research trends.

2. Research framework

2.1. Meta-learning

Meta-learning is also called learning-to-learn. Its first proposition was in the educational science community, which appeared even earlier than machine learning. The concept of “meta-learning” was first proposed by Maudsley [11] in 1979 and then was introduced into the field of machine learning. Up to now, meta-learning has become an important research branch in the field of machine learning. Many research results have already emerged because researchers are committed to using meta-learning for hyperparameter optimization, neural network optimization, and specifying the best network structure. Most existing research on meta-learning can be divided into three categories: model-based, metric-based, and optimization-based, according to different usages [12]. In addition, some other meta-learning models have been proposed in recent years. According to some research results applied in the field of cyberspace security, they can be roughly divided into two categories: online-learning-based and stacked ensemble-based methods.

Model-based methods. This type of meta-learning is based on the use of recurrent networks with external or internal memory. These technologies quickly update parameters with a minimum of training steps, and these steps can be implemented through their internal architecture or the control of other models. Meta-learning with Memory-Augmented Neural Network (MANN) [96] algorithm and meta-network are typical examples. In addition, there are Generative Adversarial Networks based (GAN-based) Meta-learning (MetaGAN) [14], Multi-response Linear Regression (MLR), Multi-Model Trees (MMT), Meta Decision Tree (MDT) [15], and many other research results.

Metric-based methods. These are techniques based on learning effective distance measurements. Their basic operational concept is like the nearest-neighbor algorithms, where their goal is to learn a measurement or a distance from objects. In addition to classical models, such as the Matching Networks (MatchingNet) proposed by Vinyals et al. [16], the Prototype Networks (ProtoNet) proposed by Snell et al. [17], the Relational Networks (RelationNet) proposed by Sung et al. [18], etc. There are other newly developed works such as convolutional Siamese neural networks, Hybrid Attention-Based Prototype Networks (HABPN) [19], Induction Networks (InductionNet) [20], SNAIL [21], Meta-SGD [22], etc.

Optimization-based methods. Through gradient descent, a small number of training samples are used to adjust the embedding model parameters. These types of meta-learning methods are all fast-learning techniques based on optimizing model parameters. Among them, the Model-Agnostic Meta-Learning (MAML) proposed by Finn et al. [23] is a relatively mature meta-learning model. It directly optimizes the learning performance of model initialization, so that even a step gradient descent from initialization can produce good results on new tasks. In addition, following the Long Short-Term Memory (LSTM) [24] technique, an LSTM-based meta-optimizer was proposed by Ravi et al. [25], which already has a lot of research and application. Besides, optimization-based meta-learning methods also include time discreteness and crawler algorithms, and so on.

Online-learning-based methods. Through the integration of online learning theory, Online Meta-learning (OML) is formed. It combines meta-learning and online learning algorithms capable of continuous lifelong learning. The study by Finn et al. [27] in 2019 shows that the OML algorithm performs better than the offline MAML. OML is, therefore, also used by some researchers to solve practical problems [28].

Stacked ensemble-based methods. Based on the traditional stacked ensemble [29] framework, the meta-learner forms a variety of meta-learner frameworks with specific advantages on different problems through the stacked integration of a variety of different machine learning algorithms. This kind of model has been used by many researchers to solve the problem of cyberspace security.

Meta-learning is a learning model different from traditional machine learning. The sample set and query set are all sampled from the labeled data set [30]. That is to say, we can construct a task set containing a large number of tasks $\Phi(train)$, such as φ_A, φ_B, and φ_C. Similarly, the task set $\Phi(test)$ contains tasks such as φ_D and is used as a set of tasks for testing. From the perspective of meta-learning, $\Phi(train)$ and $\Phi(test)$ are respectively the training set and the test set in the meta-task model F. Therefore, $\Phi(train)$ and $\Phi(test)$ can be expressed as the meta-training set and the meta-testing set, respectively. Fig. 1 depicts a schematic diagram of dataset division in meta-learning.

Generally speaking, meta-learning consists of two parts [31], one of which is to learn through internal or basic learning algorithms with internal goals to handle specific tasks well. Another learns through external or meta-learning algorithms for external targets so that meta-learners know how to adjust internal learning algorithms when given new tasks. Meta-learning is achieved by conceptually dividing learning into two levels. Through the innermost level, it can obtain exact knowledge for specific tasks, such as fine-tuning an already acquired model to consider a new data set. We include the level of cross-task knowledge across domains, such as optimizing the effective transfer between tasks. Under certain circumstances (when the outermost optimization process can perform meta-learning), automatic learning of the internal loop components can be achieved [32]. Experience is acquired through a process based on the knowledge obtained from meta-data, which is obtained from previously completed learning tasks [23]. Meta-learning technology allows rapid learning of new tasks. They use different types of meta-data, such as the properties of the learning problem, the properties of the algorithm used (for example, performance metrics), or patterns derived from data from previous problems. In this way, they can

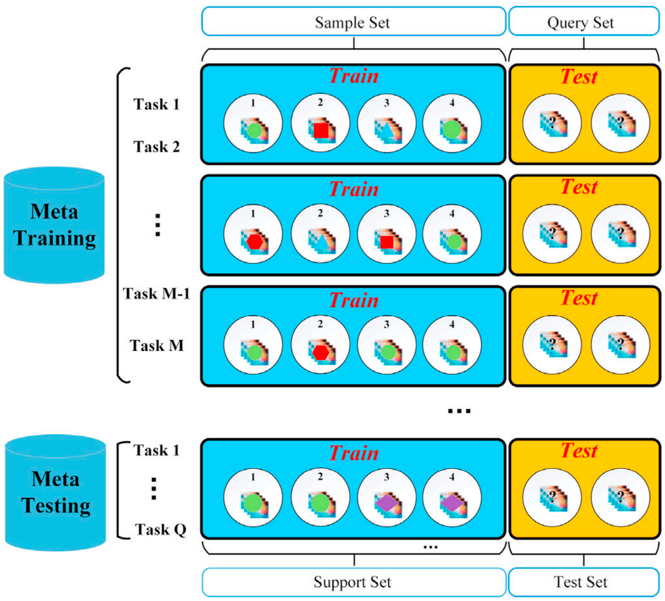

Fig. 1. Division of the datasets in meta-learning.

learn, select, change or combine different learning algorithms to effectively solve the given problem [33].

The meta-learning system needs to meet the following requirements [34]:

1. The learning subsystem must be included.
2. Experience must be derived from knowledge extracted from metadata and related to the data set under consideration or from previous learning tasks completed in similar or different fields.
3. The learning deviation should be dynamically selected.

A good meta-learning model should be trained in various learning tasks and be optimized when generalizing tasks (including tasks that may be unknown) to obtain the best performance. Each task is associated with dataset D, which contains the attribute vector and the class label of the supervised learning problem [34,35].

The best parameters of the model are:

$$\theta^{*} = \arg {}^{\min}_{\theta}\ \mathbb{E}_{D\sim P(D)}[L_{\theta}(D)] \tag{1}$$

It looks similar to the normal learning process, but in this case, the dataset is treated as a data sample. Dataset D is usually divided into training set S and prediction set B for training and testing.

$$D = \langle S, B \rangle \tag{2}$$

Dataset D contains a vector and label pairs, so

$$D = \{(X_i, y_i)\} \tag{3}$$

Each label belongs to a set of known labels L.

Defining f_θ as a classifier, parameter θ extracts the probability of data points belonging to category y, which is given by the following attribute vector x, $P_\theta(y|x)$. The optimal parameters should maximize the probability of identifying the true label in multiple training batches $B \subset D$ [33]:

$$\theta^{*} = argmax_{\theta}\mathbb{E}_{(x,y)\in D}[P_{\theta}(y|x)] \tag{4}$$

$$\theta^{*} = argmax_{\theta}\mathbb{E}_{B\subset D}\left[\sum_{(x,y)\in \mathrm{B}} P_{\theta}(y|x)\right] \tag{5}$$

Its purpose is to reduce the prediction error in the unknown labeled data, because there is a small part of "fast learning" support and its function is "fine-tuning". Fast learning is a trick that creates a "fake" data set with a small set of tags (to avoid exposing all the tags to the model). Various modifications are made in the optimization process to achieve fast learning.

A brief step-by-step description of the whole process is as follows [34,35]:

1. Create a labeled subset $L_s \subset L$.
2. Sort out a training set $L_s \subset L$ and a prediction set $B^L \subset D$. B^L and D^L are composed of data points with labels. These labels are all in set L_S, that is, $y \in L_s$, $\forall (x, y) \in S^L$, B^L.

The function of step 1 is to use the B^L in the optimization process to calculate the error and update the parameters of the model through back propagation. The implementation process step 2 is used in the simple supervised learning model. In this way, it can be assumed that each sample pair (S^L, B^L) is a data point. The model is trained so that it can be generalized with an unknown new dataset.

The following Eq. (6) is a modification of the supervised learning model, in which the symbol of the meta-learning process is added [33]:

$$\theta^{*} = argmax_{\theta}\mathbb{E}_{L_s\subset L}\left[\mathbb{E}_{S^L\subset D, B^L\subset D}\left[\sum_{(x,y)\in \mathrm{B}^L} P_{\theta}\left(x, y, S^L\right)\right]\right] \tag{6}$$

According to the general principles of meta-learning, combined with meta-learning methods constructed by using different ideas to analyze the principles used, Table 1 can be obtained.

It is worth noting that meta-learning is an algorithm that can greatly improve generalization for small samples. There are similar related research fields in machine learning, including few-shot learning [36,37], transfer learning [38], continuous learning, multi-task learning, hierarchical Bayesian models, Automatic Machine Learning (AutoML), etc. [31].

2.2. Overview

The rapid development of the Internet has exposed many problems, of which one of the most important ones is the question of cyberspace security. Cyberspace is called the fifth space, where the importance of security is not weaker than that of other spaces. Therefore, the research enthusiasm for cyberspace security has been increasing year by year. In the process of this research, a total of more than 100 papers and several monographs have been searched and read. After screening, it is found that more than 90 articles have strong relevance to meta-learning. Nearly 40 research results related to cyberspace security were further screened out. By categorizing the references in this article by year, we find that the application of meta-learning in cyberspace security has shown exponential growth in recent years, as shown in Fig. 2.

As a new field of machine learning, many excellent findings based on meta-learning have been made in recent years. This hotspot technology has also attracted the interest of researchers in other fields. Most of the research problems in these fields are strongly integrated with machine learning algorithms and models. Therefore, research on the application of meta-learning in other fields has also developed in recent years. Based on the scalability, strong generalization, and evolution of meta-learning, researchers in the field of cyberspace security have also conducted cutting-edge explorations on the combination of the two issues. This article focuses on the organic combination of cyberspace security and meta-learning, and conducts a summary study of the current status, progress and future trends in the application of meta-learning in cyberspace security.

At present, specific application research of meta-learning has been done in multiple sub-fields of cyberspace security. Based on the existing research, this article divides the meta-learning model into five research directions: model-based, metric-based, optimization-based, online-

Table 1
Meta-learning categories and their thoughts.

	Traditional classification			Online-learning based	Stack Ensemble based
	Model-based	Metric-based	Optimization-based		
Key idea	Model; Memory	Metric learning	Gradient descent	Online-learning	Stacked; Ensemble
How $P_\theta(x\|y)$is modeled?	$f_\theta(x, S)$	$\sum_{(x_i,y_i)\in S} k_\theta(x,x_i)y_i$	$P_{g_\varphi(\theta,S^L)}(y\|x)$	–	–

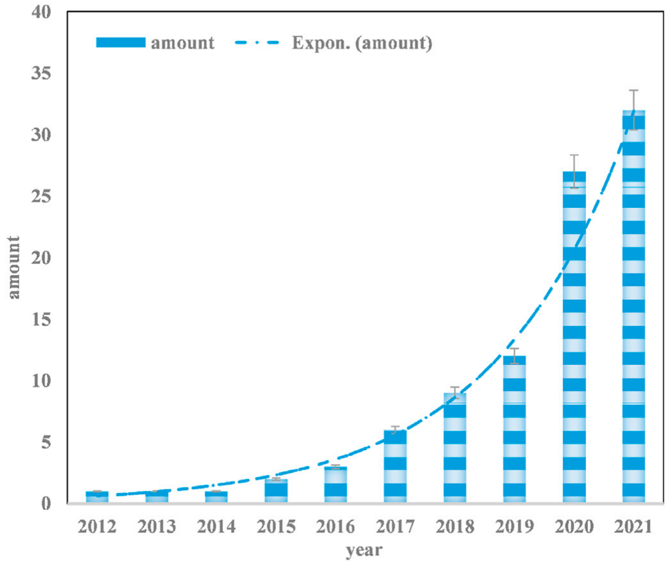

Fig. 2. The number of related research literature increased exponentially.

learning-based, and stacked ensemble-based models. Then we sort the research progress and core ideas in these subdivisions. Moreover, we adopt a new classification method to divide cyberspace security issues into three major branches from a macro perspective: cyber security, information security, and artificial intelligence security. Then we subdivide each branch, classify and summarize the application research of the corresponding meta-learning model in these sub-fields and then form a new research framework, as shown in Fig. 3.

3. Meta-learning applied to cyber security

3.1. Cyber source management

Cyber source management refers to the management of various resources in computer networks, such as computing power, electric power, etc. In recent years, most of the research in this field has been based on traditional machine learning methods. However, due to its large data demand and resource occupation, relevant researchers have been exploring new solutions. The application of meta-learning in Mobile Edge Computing (MEC) [39–41], power forecasting, and smart grid is summarized in this subsection.

MEC is a new architecture which uses a mobile base station to extend cloud computing services to the network edge [42]. As an important component of the 5G architecture, it supports various innovative applications and services requiring ultra-low latency [43]. In 2020, Huang et al. [44] proposed an MEta-Learning-based computation Offloading (MELO) algorithm for dynamic computing tasks in MEC networks. In order to solve the problem of Deep Neural Network (DNN), a meta-learning model based on MAML is introduced [23]. Specifically, it trains a general DNN for different types of MEC task scenarios and can quickly learn to adapt to new MEC task scenarios after several training iterations [45–47]. The numerical results show that the accuracy of the

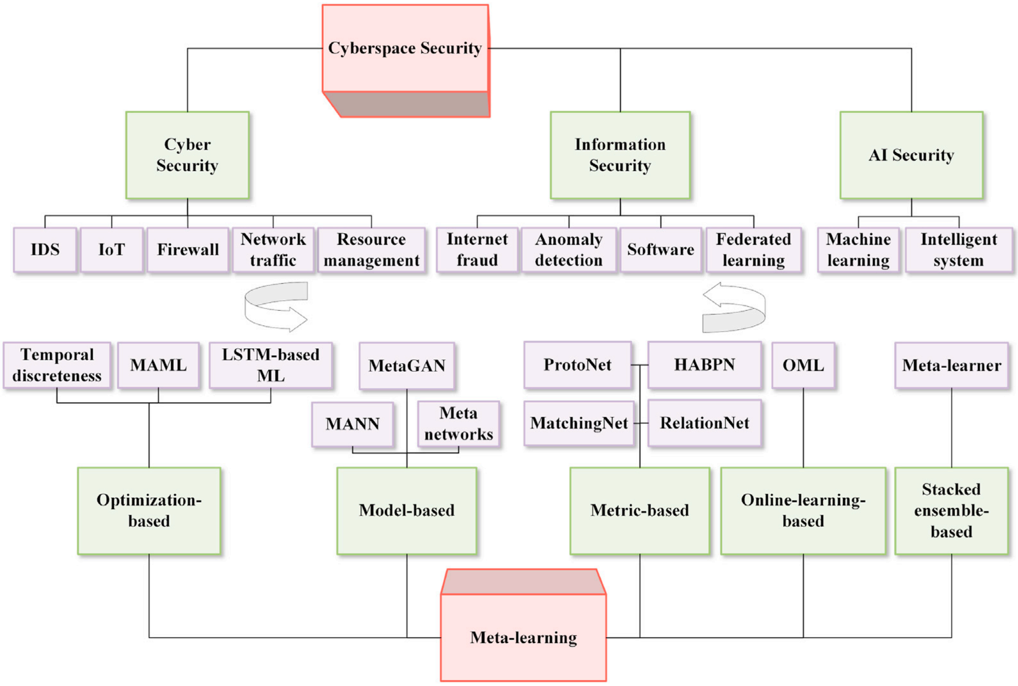

Fig. 3. Research framework of meta-learning applied to cyberspace security.

MELO algorithm is more than 99% after one-step fine-tuning of 10 training samples.

The security of electric power systems provides the most basic guarantee for the security of computer networks, so the research on the smart grid and electric power forecasting is closely related to cyber security [48,49]. In 2020, Li and Hu [50] proposed a novel OML framework which uses an LSTM-based meta-optimizer to solve the adaptive problem of the smart grid continuous model. The algorithm continuously adjusts the pre-trained basic learners by effectively extracting the real-time data and then uses the meta-optimizer to adaptively control the parameters of the basic learners. The simulation results show that the performance of meta-learning and OML is better than that of online basic learning. At the same time, when the training data is limited or the training and real-time data have very different time-varying patterns, the performance of OML is better than meta-learning and online basic learning. In 2021, Lee and Rhee [32] adopted the two methods of transfer learning and meta-learning, which can be smoothly integrated into the deep neural network to predict the power load. This method is based on MAML, and it is mainly committed to developing a generalized learning model with few samples so as to solve the problem that the individual model is difficult to train and needs a lot of time to accumulate personalized data. Then the proposed method is tested on residential and non-residential datasets. Compared with traditional methods, the RMSE improvement rates of the meta-learning model are 7.84% and 15.07%, respectively. In the same year, for battery safety monitoring, Ding et al. [51] proposed an effective prediction method for early warning of Thermal Runaway (TR), and named it Meta-TRFNN. The algorithm has a novel data-driven method, which can accurately predict the TR state of cells in advance. In the algorithm, meta-learning framework is used to deal with the problem of insufficient data. Algorithms based on LSTM, Gated Recurrent Unit (GRU) [52] and Recurrent Neural Network (RNN) [53] are stacked and integrated. Then they evaluated the Meta-TRFNN on simulated samples and proved its predictive ability. They also told the benefits of involving high-dimensional thermal images and the effectiveness of the meta-learning framework. The stacked ensemble-based meta-learning has obvious advantages in solving problems in some areas. Fig. 4 is its schematic diagram.

3.2. Intrusion detection

Intrusion Detection System (IDS) [54] is the key research direction of current cyber security research. IDS is used to effectively detect all kinds of malicious attacks on the cyber and is one of the key systems to maintain cyber security. Although the machine learning model is widely used in intrusion detection and has made great progress in solving related problems, the traditional machine learning needs a large number of samples for training, which leads to some bottlenecks in this field. For example, the zero-day attack [55] is an attack launched on the day of vulnerability discovery. It is difficult for security agencies to obtain sufficient attack samples in a short time. Therefore, it may be too late to build and publish data sets.

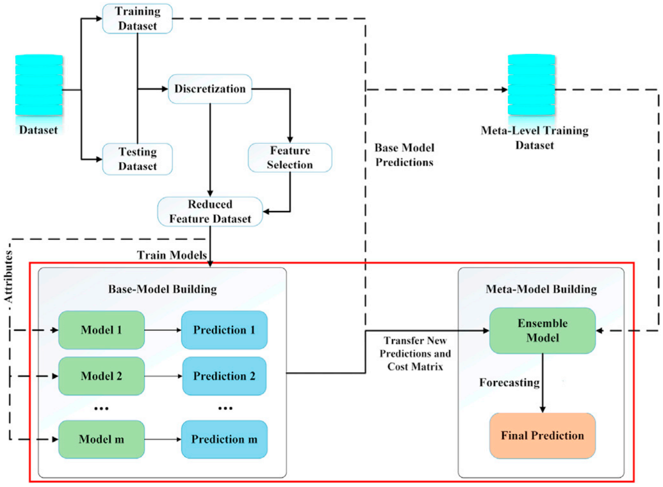

Fig. 4. General meta-learning model based on stacked ensemble.

Detection based on a limited number of attack samples is a few-shot intrusion detection problem [30]. In 2020, Olasehinde [15] first applied the idea of meta-learning to the solution of intrusion detection and proposed a new IDS method based on three meta-level algorithms. This method uses the stacked ensemble framework to build IDS and uses Naive Bayes (NB) and Decision Tree (DT) to train three meta-learning algorithms: MDT, MLR, and MMT. Then, it uses the UNSWNB15 test data to evaluate the basic layer model and the meta stack model. It shows that the intrusion detection accuracy of the three meta-learner models is higher than the best accuracy of each original basic model. In the same year, chalé et al. [56] proposed another intrusion detection framework based on this kind of meta-learning idea, which combines the user input and data element characteristics to select the best algorithm to detect cyber-attacks. Then experiments are carried out on the NSL-KDD dataset. It was found that the framework eliminates the guesswork of common trial and error algorithm selection methods and selects the algorithm with higher classification performance. Xu et al. [30] applied the mature meta-learning theory to IDS's few-shot detection for the first time in 2020 and proposed a detection method FC-net based on a meta-learning framework. The algorithm is based on the principle of a deep neural network, which is mainly composed of a feature extraction network and a comparison network. Experiments show that the proposed network intrusion detection method is universal and not limited to specific types of attacks. The results of training and testing on data sets show that the proposed method can achieve an average detection rate of 98.88%. In a few cases, malicious samples in the untrained dataset can be successfully detected, with an average detection rate of 99.62%.

Then in 2021, Ahsan et al. [57] adopted the idea of stacked ensemble meta-learning based on the Dynamic Feature Selector (DFS), and integrated CNN + LSTM, Bi-directional LSTM (BiLSTM) [53], GRU, DT [58] and random forest algorithms. According to the training results of each instance, it is important to dynamically determine which features to generate high prediction accuracy. Experimental results show that for NSL-KDD, the algorithm reduces the feature size of a hot-coded feature from 123 to 50 and improves the accuracy from 99.54% to 99.64%. In UNSWNB15, the accuracy is improved from 90.98% to 92.46%, and the feature size is reduced from 196 to 47. The proposed method can significantly reduce the number of features required for processing and obtain a higher accuracy at the same time.

3.3. Internet of things security

Internet of Things (IoT) is a network of computer-embedded physical devices operating in the cyber or digital space [59]. One of its key network characteristics is that they can communicate with each other, exchange information, and send/receive instructions in wired or wireless communication channels. Like those technologies that support the computer network or connect to the Internet, the rise of IoT will also be accompanied by large-scale cyber security attackers who use the IoT to commit network crimes, making network security unable to be guaranteed [60].

To quickly identify malicious IoT devices, network behavior fingerprint inference has become urgent. In 2020, Pan [61] applied the metric-based meta-learning algorithm to solve the problem and developed a model called DeepNetPrint, which solved the problem of inferring the fingerprint of IoT network behavior with the limited traces of network activities. The model combines an automatic encoder based on the Convolutional LSTM (ConvLSTM) [62] and a kind of time series ProtoNet. Compared with other similar models, it is found that the new model performs better in identifying IoT devices, even if the new devices have not been trained. Another problem in the field of IoT is sensor recognition. Aiming at the multi-label classification problem, Lin et al. [63] proposed a new multi-label classification meta-learning-based model. It works on the integration of clustering algorithms and the Generalized Linear Mixed Model (GLMM). In this algorithm, the clustering stage is used to capture the association between tags while reducing the computational complexity of a large number of tag subsets. Experimental results show that the performance of the algorithm is better than other algorithms, especially in the case of a large number of tags. Then, in 2021, Mishra et al. [64] made a new breakthrough. They developed a stacked ensemble meta-learning model for various cyber security threats in the IoT. The model integrates five algorithms: DT, K-Nearest Neighbor (KNN), BT, Stochastic Gradient Descent (SGD), and Support Vector Machines (SVM), which can enhance the performance of the basic machine learning model for anomaly detection in IoT devices. The proposed model learns from the prediction error of the basic classifier, and a more accurate prediction model is established.

3.4. *Other fields*

In the past two years, the application of meta-learning has shown explosive growth. For cyber security problems, in addition to cyber source management, intrusion detection, and IoT, there are also preliminary explorations on the application of meta-learning in other fields, such as spam detection, phishing attacks [65], and network traffic prediction. In 2012, Chen [8] proposed a meta-learning string kernel Core Vector Machines (CVM) to satisfy the string format stream data learning and applied it to spam detection. Experiments show that the method can outperform the traditional string kernel method in accuracy and efficiency. The advantage of this method is that the defined string meta-features can be used to extract meta-knowledge from any string dataset, and the process of applying meta-learning to string classification and using string meta-features by extracting meta-knowledge is explained.

Phishing attack is the most famous behavior of cheating Internet users, in which the actor acts as a trusted entity. It is realized by abusing the inadequate protection provided by electronic tools, making use of the ignorance of user objects, and illegally obtaining personal data, such as sensitive private information and passwords. In 2020, Zhu [66] proposed an online meta-learning firewall to prevent phishing attacks. It is a highly innovative and fully automated active security tool which uses the LSTM meta-learner algorithm. This method can learn to use a small number of samples for effective classification. At the same time, it can converge in very few steps. Inspired by the LSTM model, the proposed system is an improvement of KNN's self-adjusting memory algorithm. The system can implement decision rules best, classify and detect phishing attacks, and enhance the active security of the digital system.

With the development of intelligent devices, the network traffic increases rapidly, which makes network traffic analysis and prediction modeling more and more important. In addition, to achieve fast and secure communication, many fields, including congestion control and network management, need to accurately predict network traffic [67]. Kim [68] proposed MetaCGAN in 2021 based on MetaGAN proposed by Zhang et al. [14]. It uses Conditional Generative Adversarial Networks (CGAN) and data enhancement based meta-learning methods to overcome insufficient network traffic data in attack analysis. Experimental results show that the proposed technique can overcome the problem of data shortage and outperform Multi-Layer Perceptron (MLP) regression.

3.5. *Summary*

Through the above research, we find that there are many research results of applying meta-learning to solve the cyber security problem. According to the characteristics of each subdivided problem domain, researchers have adopted different meta-learning strategies and verified their advantages and accuracy through experiments, as shown in Table 2. At present, the meta-learning model has achieved better results in MEC, IDS, spam detection, IoT anomaly detection, phishing attack detection, and other problems that traditional machine learning is good at solving. As a mature meta-learning model, MAML has strong robustness in application. Therefore, its application in other problem areas can be further explored. It is also found that the combination model of OML and LSTM-based meta-optimizer has strong adaptability in solving practical problems, and it is worth further exploring its next optimization direction. In addition, the stacked ensemble-based meta-learning uses a variety of algorithms to stack, and the integrated model has outstanding advantages in solving classification problems. It also can be used in fields such as grinding, which could support the industry-4.0 [69]. Therefore, the meta-learning model has a great prospect in dealing with various problems and is worth promoting in the future.

4. Meta-learning applied to information security

4.1. *Information system security*

There are many kinds of information systems, most of which are database-based. These information systems usually store data that is vital to individuals, enterprises, and countries. Once these information systems have security problems, the impact on the current society is extremely bad. This section focuses on the application of meta-learning in information systems, including electronic payment system security, system anomaly detection, and software system security.

The electronic payment system has always been an important tool for business transactions all over the world, in which the credit card has become a means of payment. At present, many traditional machine learning technologies and methods have been proposed to mitigate this common threat. However, recently, meta-learning methods have been introduced and have achieved some results. Pun [70] first proposed to use a meta-classifier as the filter of credit card transaction data. The meta-classifier uses the predictions obtained from the three basic classifiers to get the final prediction of any transaction. The model integrates NB, DT, and KNN to construct three basic classifiers. The NB algorithm is also used to combine basic classifier predictions at the meta-level to construct the final classifier. Subsequently, in 2020, Olowookere and Adewale [71] proposed a fraud detection framework combining the stacked ensemble-based meta-learning and cost-sensitive learning paradigm. The framework integrates KNN, DT, and MLP algorithms, which allows the basic classifiers to carry out the traditional matching. In the process of ensemble learning, cost-sensitive learning is introduced to match the cost-sensitive meta-classifiers, so it is not necessary to carry out cost-sensitive learning for each basic classifier. The results of invisible data classification show that the cost-sensitive ensemble classifier maintains a good AUC value, which shows that it has consistent performance in datasets with different fraud rates.

The purpose of anomaly detection is to identify data objects or behaviors that are significantly different from most anomalies. Anomaly detection has important applications in fraud detection, network security attack detection, medical diagnosis, and other fields [72]. Many anomaly detection algorithms have been proposed, but they are usually unsupervised hypothetical anomaly patterns [73,74]. In order to deal with the challenge of the high false positive rate of traditional anomaly detection algorithms, Zha et al. [7] proposed a Meta-learning model of model-based Active Anomaly Detection (Meta-AAD), which is a new framework of learning meta-strategy for query selection. The model uses deep reinforcement learning to train meta-strategy to select the most appropriate instance. A large number of experiments show that Meta-AAD is significantly better than the latest reordering strategy and unsupervised method, and the trained meta-strategy is transferable and easy to deploy.

Malware is a kind of malicious attack which aims to penetrate the security, integrity, and functionality of the system. Different types of malware include viruses, worms, trojans, backdoors, spyware, botnets,

Table 2
Application of meta-learning in cyber security.

Fields	Feature	Meta-learning model	Advantage	Related work
MEC	The learning speed is slow; the sample size required to train the model under dynamic characteristics is large; it takes up a lot of resources	MAML	Quickly learn to adapt to new MEC task scenarios	[44]
Smart grid	Continuous model adaptation	OML; LSTM-based meta-optimizer	Better than meta-learning and online basic learning	[50]
Electric forecasting	Individualized models are difficult to train; it takes a lot of time to accumulate personalized data	MAML	Few-shot learning with strong generalization; an effective tool for short-term load forecasting tasks	[32]
Battery TR forecasting	Large amount of data required	Stacked ensemble-based	Novel data-driven approach; make accurate multi-step forecasts ahead of time	[51]
IDS	Need a large sample set; unable to stop zero-day attacks	Model-beased; Stacked ensemble-based	High accuracy of intrusion detection; not limited to specific types of attacks	[15,30,56,57]
IoT behavioral fingerprint	Limited traces of network activity	Metric-beased	Relatively good performance in identifying IoT devices (new devices)	[61]
Multi-label classification	Multi-category	Stacked ensemble-based	Reduce computational complexity; Low dependence; stable performance	[63]
IoT anomaly detection	Multiple threat detection	Stacked ensemble-based	Enhance the performance of basic machine learning models; learn from prediction errors	[64]
Spam detection	Learning of stream data in string format	Model-beased	It can outperform traditional string kernel methods in terms of accuracy and efficiency; meta-knowledge can be extracted from any string data set	[8]
Firewall; Phishing attack	High automation; small sample size	OML; LSTM-based meta-optimizer	Few convergence steps; use a small number of samples for effective classification; strong active security	[66]
Network traffic analysis	Large amount of network traffic data; low sample size available	MetaGAN	Overcome the lack of network traffic data in attack analysis	[68]

etc. With the popularity of Internet users, malicious software poses a serious threat to the security of computer systems [5]. In the aspect of software system security, finding and repairing security vulnerabilities in advance is one of the effective methods to prevent malicious activities. Aiming at malware attacks, Tran et al. [75] first applied Neural Turing Machine (NTM) in 2018 to classify malware sets with only a small number of known samples by using the MANN algorithm. MANN first slowly learns these malware features in a domain. In the other domain, two trained networks (Embedded Network and MANN) are used to classify malware sets that have never been seen before. Experiments show that using the one-shot/few-shot algorithm can solve the problem of rare malware detection and classification. Different from Ref. [75], Kang and Kim [76] believed that the binary file of each malware is a gray image. They used MANN to receive a series of input images and evaluated the current input according to its memory and the previous hidden state of the controller. Then, based on the work of [76], in 2019, Tran et al. [77] further proposed a method to solve the problem of malware classification by using a few-shot learning algorithm and fusing two classical models of MatchingNet and ProtoNet. This method solves the problem of a small sample size to a certain extent and achieves higher accuracy than traditional machine learning in experiments. Then, in 2020, Tang et al. [78] proposed a new neural network structure called ConvProtoNet based on ProtoNet.

This structure embedded meta-learning models such as relational network, HABPN, induction network, SNAIL, and Meta-SGD on the basis of the work of Trung Kien Tran et al. to make the classifier robust enough to adapt to the unknown malware categories. Even in extreme cases, with only 5 samples per category, ConvProtoNet can achieve an average accuracy of more than 70%, and its performance is better than other traditional malware classification methods in experiments on multiple data sets. ConvProtoNet is divided into three parts: the embedding module which projects the data set into the feature space f, the convolution induction module which generates the prototype g, and the softmax classifier generator which uses the modified cosine similarity function C, as shown in Fig. 5.

For the problem of security vulnerabilities, because of their complexity, the discovery of security vulnerabilities highly depends on human experts. Therefore, as one of the ways to find bugs, the source code audit costs a lot, and the quality of audit varies from performer to performer. In recent years, the machine learning technology has developed rapidly and its accuracy in many tasks is higher than that of humans. However, in most cases, it is very difficult to obtain legal training data, and the less data often means the more fatal the security problem. In order to overcome these obstacles, in 2020, Shin [79] proposed a deep neural network model for discovering security vulnerabilities based on meta-learning. The language model uses the converter based on sub-word tagging and self-attention in natural language processing to realize the understanding of source code and uses the mature model MAML to overcome the lack of reliable vulnerability samples in the deep learning model. Through experiments, it is found that the model can identify DOM-based XSS errors which are difficult to find by traditional detection methods. The evaluation results showed that the model was 45% higher than the baseline in the F1 score.

4.2. Information content security

Information security is the core of cyberspace security. It can be said that the goal of all security issues on the Internet is to obtain or modify the damaged data information content. Federated learning [80–82], as a new theory proposed in recent years, aims to protect the data privacy of users by maintaining the shared model on the central server and using the data of all clients to train the model cooperatively in a distributed way. In many scenarios, such as smart medicine, data are distributed among different clients and sensitive to privacy, so it is unrealistic to collect the original data from the central server for model training. Chen et al. [83] applied meta-learning to federated learning for the first time in 2019 and proposed a federated meta-learning framework FedMeta, which used MAML to perform k-order classifier initialization meta training for all n categories. Compared with FedAvg, which is the leading optimization algorithm in federated learning, its accuracy rate is improved by 3.23%–14.84%. In addition, only the parameterized algorithm is transmitted between the mobile device and the central server under the framework, and no original data is collected from the server, thus preserving the user's privacy to the maximum extent. Then in 2020, Lin et al. [84] proposed a federal meta-learning algorithm based on Distributed Robust Optimization (DRO) and proved its convergence under certain conditions. It overcomes the vulnerability of MAML to adversarial attacks. Experiments on different datasets show the effectiveness of the federated

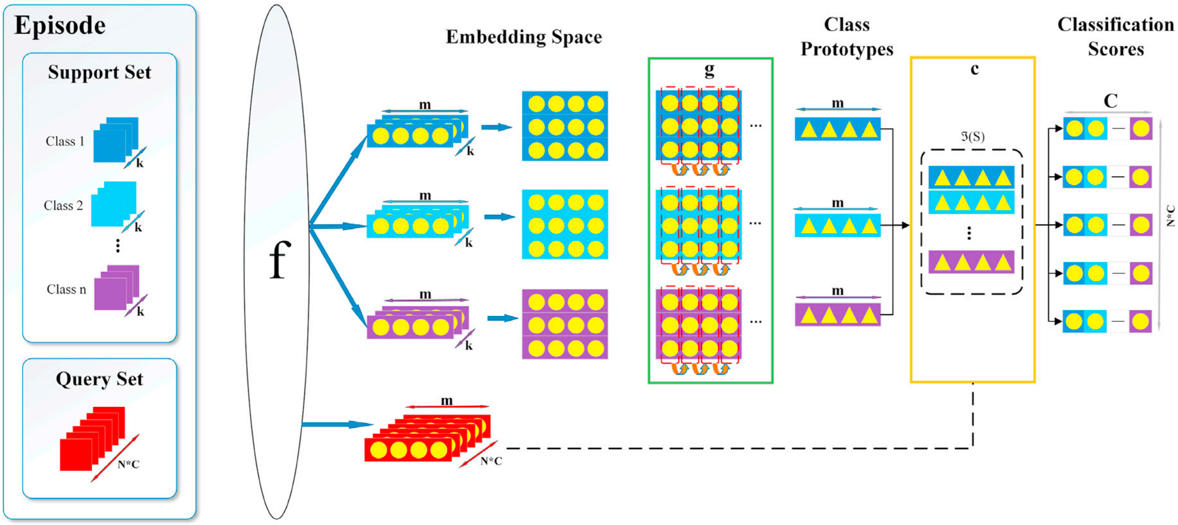

Fig. 5. ConvProtoNet meta-learning model based on metric. Adapted from Ref. [78].

meta-learning framework, which contributes to the research of edge intelligence.

4.3. Summary

Information security is a major branch of cyberspace security. The meta-learning model has been applied to various problems in information and information system security. It can be seen from Table 3 that except OML, which is not applied in this field, the other four kinds of meta-learning models have practical applications in the given problem. The current research results show that, firstly, although the meta-learning model based on stack ensemble can solve the classification problems such as fraud detection, there is still room for further improvement due to its high time cost. Secondly, the meta-learning models, by extracting learning meta-knowledge, fit the problem of sensitive information. For example, to protect data privacy, the combination of federated learning and MAML achieves good experimental results. In addition, there are many studies on meta-learning in malware detection, including model-based, metric-based and stacked ensemble-based meta-learning models. We can continue to explore the application effect of the other two categories of meta-learning models. Anomaly detection and malicious code audit are two important problem areas. In addition, in view of the current trend of industrial intelligence, data visualization analysis in the field of metallurgy can also be used as the next step in the application research of meta-learning [85]. According to their need for high intelligence, we should continue to study the application of meta-learning in them.

5. Meta-learning applied to artificial intelligence security

5.1. Machine learning algorithm security

Adversarial attacks against machine learning models have been studied in the field of machine learning and security, as well as in many different model types [86]. The latest work of adversarial machine learning shows that many deep neural networks are vulnerable to the attack of adversarial examples, which is the disturbance input designed by the attacker to make the model wrong [87]. Due to the excellent performance of meta-learning, there are applications of meta-learning in the research of adversarial attacks. In 2017, Muñoz-González et al. [88] proved that meta-learning can be used to create training time attacks against simple linear classification models. However, for the continuous data, this study is rarely successful in attacking deep neural networks, while for the discrete datasets, the deep learning model for more than two classes is not considered. In contrast to the work of Zügner et al. [89], Zügner and Günnemann [90] proposed an algorithm for the global attack on the (depth) node classification model based on the principle of meta-learning in the field of adversarial attack on Graph Neural Networks (GNN) in 2019. Using meta-gradients can solve the two-level optimization problem of the poisoning attack. Experimental results show that the attacks generated by this method will lead to a significant decline in the classification performance of the GNN model and even transfer to the unsupervised model. In 2020, Zhou et al. [91] proposed a Robust Meta Network Embedding (ROMNE) framework based on MAML for the graph mining model which is vulnerable to adversarial attacks. The framework uses the meta-learning mechanism to learn adaptive parameters, which improves the robustness of multiple network embedding on adversarial noisy networks. At the same time, it retains the utility in the original single network. Fig. 6 shows an example of a metric-based meta-learning model MAML embedded in a powerful meta network with two analysis components.

In addition, Edmunds et al. [92] conducted a study on the adversarial attack of MAML and proposed a series of experiments to test the sensitivity of MAML to malicious attacks. After the experiment, through the steps taken on the vulnerability of transfer attacks based on the MAML model, the defense method against this kind of attack may be improved. Subsequently, Yin et al. [13] also proved that the performance of MAML in antagonistic samples decreased significantly. In order to improve the robustness of meta-learning, they proposed a meta-learning algorithm called ADML, which uses clean samples and hostile samples to promote internal gradient updates to fight against meta updates. The disadvantage is that the structure is too complex to achieve.

5.2. Intelligent system security

With further development of the next-generation Internet and 5G technology [93], intelligent systems will be more inseparable from our lives. Therefore, research on the security of intelligent systems will be of great significance. The following is an introduction to the preliminary application of meta-learning in the intelligent recommendation system and the intelligent security system.

Intelligent recommendation system will recommend items to users according to the past behaviors of other users (the goods they have purchased, browsed, or selected and their evaluation of these goods). This kind of system will effectively recommend products or projects that users may be interested in. However, due to the dependence of the

Table 3
Application of meta-learning in information security.

Fields	Feature	Meta-learning model	Advantage	Related work
Fraud detection	High security level; difficult to detect	Stacked ensemble-based	High reliability; high accuracy	[71]
Anomaly detection	High false alarm rate; poor initiative	Model-based	Is transferable; easy to deploy	[7]
Malware classification	Threatening; small sample size	MANN; Metric-based; Stacked ensemble-based	Can classify rare malware; solved the problem of small sample size; superior to other traditional malware classification methods	[75–78]
Source code auditing	Prone to false positives and omissions; lack of legal training data	MAML	Overcome the lack of reliable vulnerability samples; high recognition accuracy	[79]
Federated learning	Training across multiple nodes; data sensitive	MAML	Only parameterized algorithms are transmitted between nodes; significantly improved accuracy	[83,84]

collaborative filtering technology on external information sources, it is vulnerable to profile the injection attack, which is called the shilling attack. To solve this problem, Bhebe and Kogeda [94] proposed a new framework based on meta-learning in 2015. The framework is based on stacked ensemble learning and integrates NB, SVM, and KNN. Then it uses the meta classifier to detect the shilling attack. Experiments show that the comprehensive performance of the proposed meta-learning classifier is 99%, which is better than the neural network and the nearest neighbor classifier.

Most Critical Infrastructure Network (CIN) is related to energy, communication, information technology, transportation, national defense, government facilities, and industry. Therefore, the intelligent safety system, such as the newest intelligent vehicle systems, plays an important role in maintaining important social functions and protecting critical infrastructure. However, the security system based on the traditional machine learning algorithm needs continuous training and cannot prevent advanced zero-day attacks. Therefore, in 2020, Bing [33] proposed a model-based external storage meta-learning prototype compatible with the NTM and took it as the method to build the model. The introduction of the MANN model allows the memory of useful data from past processes by integrating external memory. The efficiency of the system is successfully tested in an extremely complex scene, which greatly improves the detection efficiency. Dibaei et al. [95] investigated attacks on intelligent vehicle systems and showed that the available defences against these attacks are cryptography, network security, software vulnerability detection, and malware detection.

5.3. Summary

It should be noted that artificial intelligence security will be the core security issue in the future intelligent era, which directly determines the reliable landing of various intelligent achievements and devices in the future. According to the existing research results (Table 4), it can be found that MAML is used in other machine learning algorithms to resist attacks, and its security is also considered by some frontier scholars. The research on the security of intelligent systems also has strong practical significance. Remaining challenges and future directions for prevention have been discussed as well. Next, we can try to apply MAML to the security defense of the CIN system and the detection of shilling attacks.

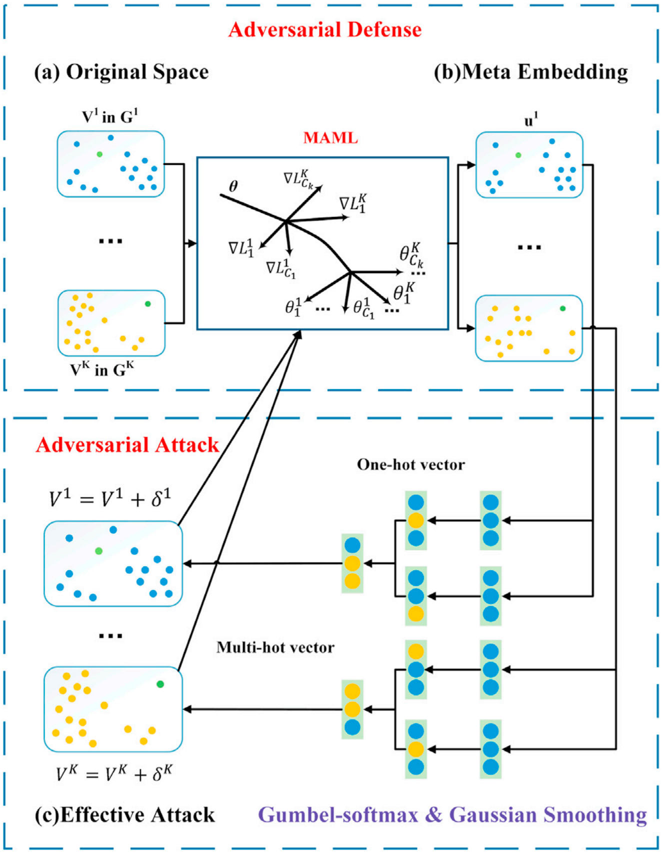

Fig. 6. MAML-based ROMNE framework. Adapted from Ref. [91].

Table 4
Application of meta-learning in AI security.

Fields	Feature	Meta-learning model	Advantage	Related work
Adversarial attack in GNN	Disturbance input; the attack algorithm itself	MAML	Solved the second-tier optimization problem of poisoning attacks; global attack; high robustness; adaptable	[11, 88–90]
Adversarial attack in MAML	Use clean samples and hostile samples to promote internal gradient updates to fight against meta-updates; exists adversarial attack vulnerability	–	–	[13,91]
Shilling Attack Detection	Rely on external information sources; profile injection attack	Stacked ensemble-based	Comprehensive performance reaches 99%; better than neural network and KNN	[93]
CIN	Continuous training; unable to stop zero-day attacks	MANN	Allows to remember useful data of past processes; high detection efficiency	[33]

6. Conclusion

This paper focuses on the current problems in the field of cyberspace security and mainly reviews the application research results of meta-learning in this field. This paper adopts the methods of investigation, experiment, and literature research. By investigating the research content and bottleneck problems in the field of cyberspace security, this paper explores the latest model of meta-learning in the field of machine learning. On this basis, through principle analysis, comparative experiments, literature comparison, and extensive investigation, a new classification method of meta-learning and cyberspace security is obtained. The paper firstly summarizes the current research status of meta-learning and its application in cyberspace security and proposes a new research framework for meta-learning in cyberspace security. Then, according to the framework, the application research results of the current five research directions of meta-learning in the three branches of cyberspace security are analyzed. And the advantages and disadvantages of meta-learning in different application scenarios and a comparison of the proposed framework with traditional machine learning are analyzed. Also, according to the current research status, the future research direction of each branch is proposed.

Furthermore, our systematic investigation and research show that although meta-learning has a great advantage in dealing with the problem of a few samples (which is very useful for cyberspace security), there are still some shortcomings which bring about some negative impacts on the development of meta-learning. The biggest characteristic of model-based meta-learning is its flexibility, but it is accompanied by the problem of weak generalization. The advantage of the metric-based meta-learning is its simplicity and efficiency, but it is limited to supervised learning. The optimization-based meta-learning has a more robust generalization but is more computationally expensive. The online-learning-based meta-learning has good synchronization ability and can carry out continuous learning, but it has high complexity and requires many computer resources. The advantage of the stacked ensemble-based meta-learning is its flexibility and reliability, and its disadvantage is the high computation and time costs. In addition to the cyberspace security field, meta-learning has been applied in artificial intelligence, simulation, natural language processing, metallurgy, medicine, and other fields.

In the future, the main challenges of meta-learning are generalization ability, computational complexity, and task migration. For the multi-distribution case of the meta-training set, if it cannot match the corresponding distribution of the meta-testing set, meta-learning will not be able to proceed. This feature makes its generalization ability weak, which needs to be further studied. Meta-learning usually involves bi-level optimization, which makes its calculation very complex. How to improve the efficiency of calculation is a difficult problem. In addition, the training mode of meta-learning is multitasking, and only specific task families can be used in specific tasks, which greatly limits the representation and dissemination of knowledge. Meta-learning can play an excellent role in many fields (especially in solving problems with less data). Researchers who are interested in this field can further improve meta-learning in view of the above challenges.

Finally, meta-learning solves the bottleneck problem of cyberspace security to a certain extent, and it is the next research object for researchers. Meta-learning is a novel model which is different from the traditional training model of machine learning. It divides the data into multiple tasks and overcomes the problem of few-shot. At the same time, combined with the traditional machine learning model for feature extraction and parameter optimization, a new model with self-learning ability is obtained, which is a feasible way to strong AI in the future.

Declaration of competing interest

The manuscript has not been published before and is not being considered for publication elsewhere. All authors have contributed to the creation of this manuscript for important intellectual content and read and approved the final manuscript. We declare there is no conflict of interest.

Acknowledgements

This work was supported by Hebei Province Natural Science Fund for Distinguished Young Scholars (NO. E2020209082).

References

[1] J. Luo, M. Yang, Z. Ling, W. Wu, X. Gu, Architecture and key technologies of cyberspace security, Scientia Sinica Inform. 46 (8) (2016) 939–968.

[2] A.L. Buczak, E. Guven, A survey of data mining and machine learning methods for cyber security intrusion detection, IEEE Commun. Surv. Tutor. 18 (2) (2015) 1153–1176.

[3] I.H. Sarker, A. Kayes, S. Badsha, H. Alqahtani, P. Watters, A. Ng, Cybersecurity data science: an overview from machine learning perspective, J. Big Data 7 (1) (2020) 1–29.

[4] Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou, C. Wang, Machine learning and deep learning methods for cybersecurity, IEEE Access 6 (2018) 35365–35381.

[5] S. Mahdavifar, A.A. Ghorbani, Application of deep learning to cybersecurity: a survey, Neurocomputing 347 (2019) 149–176.

[6] Y. Pan, Y. Wang, P. Zhou, Y. Yan, D. Guo, Activation functions selection for bp neural network model of ground surface roughness, J. Intell. Manuf. 31 (8) (2020) 1825–1836.

[7] D. Zha, K.-H. Lai, M. Wan, X. Hu, Meta-aad: active anomaly detection with deep reinforcement learning, in: IEEE International Conference on Data Mining (ICDM), IEEE, 2020, pp. 771–780.

[8] Y.G. Chen, On-line Fast Kernel Based Methods for Classification over Stream Data (With Case Studies for Cyber-Security), Ph.D. thesis, Auckland University of Technology, 2012.

[9] J. Vanschoren, Meta-learning: A Survey, 2018 arXiv:1810.03548.

[10] P. Brazdil, C.G. Carrier, C. Soares, R. Vilalta, Metalearning: Applications to Data Mining, Springer Science & Business Media, 2008.

[11] D.B. Maudsley, A Theory of Meta-Learning and Principles of Facilitation: an Organismic Perspective, 1980.

[12] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, T. Lillicrap, Meta-learning with memory-augmented neural networks, in: International Conference on Machine Learning, PMLR, 2016, pp. 1842–1850.

[13] C. Yin, J. Tang, Z. Xu, Y. Wang, Adversarial Meta-Learning, 2020 arXiv: 1806.03316.

[14] R. Zhang, T. Che, Z. Ghahramani, Y. Bengio, Y. Song, Metagan: an adversarial approach to few-shot learning, in: Proceedings of the 32nd International Conference on Neural Information Processing Systems, 2018, pp. 2371–2380.

[15] O.O. Olasehinde, O.V. Johnson, O.C. Olayemi, Evaluation of selected meta learning algorithms for the prediction improvement of network intrusion detection system, in: IEEE International Conference in Mathematics, Computer Engineering and Computer Science (ICMCECS), IEEE, 2020, pp. 1–7.

[16] O. Vinyals, C. Blundell, T. Lillicrap, D. Wierstra, et al., Matching networks for one shot learning, Adv. Neural Inf. Process. Syst. 29 (2016) 3630–3638.

[17] J. Snell, K. Swersky, R. Zemel, Prototypical networks for few-shot learning, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 4080–4090.

[18] F. Sung, Y. Yang, L. Zhang, T. Xiang, P.H. Torr, T.M. Hospedales, Learning to compare: relation network for few-shot learning, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2018, pp. 1199–1208.

[19] T. Gao, X. Han, Z. Liu, M. Sun, Hybrid attention-based prototypical networks for noisy few-shot relation classification, in: Proceedings of the AAAI Conference on Artificial Intelligence vol. 33, 2019, pp. 6407–6414.

[20] R. Geng, B. Li, Y. Li, Y. Ye, P. Jian, J. Sun, Few-shot Text Classification with Induction Network, 2019 arXiv:1902.10482.

[21] N. Mishra, M. Rohaninejad, X. Chen, P. Abbeel, A simple neural attentive meta-learner, in: International Conference on Learning Representations, ICLR), 2018.

[22] Z. Li, F. Zhou, F. Chen, H. Li, Meta-sgd: Learning to Learn Quickly for Few-Shot Learning, 2017 arXiv:1707.09835.

[23] C. Finn, P. Abbeel, S. Levine, Model-agnostic meta-learning for fast adaptation of deep networks, in: International Conference on Machine Learning, PMLR, 2017, pp. 1126–1135.

[24] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.

[25] S. Ravi, H. Larochelle, Optimization as a model for few-shot learning, in: International Conference on Learning Representations (ICLR), 2017.

[27] C. Finn, A. Rajeswaran, S. Kakade, S. Levine, Online meta-learning, in: International Conference on Machine Learning, PMLR, 2019, pp. 1920–1930.

[28] D.A.E. Acar, R. Zhu, V. Saligrama, Memory efficient online meta learning, in: International Conference on Machine Learning, PMLR, 2021, pp. 32–42.

[29] T. Dietterich, et al., Ensemble learning, in: The Handbook of Brain Theory and Neural Networks, MIT Press, 2002.

[30] C. Xu, J. Shen, X. Du, A method of few-shot network intrusion detection based on meta-learning framework, IEEE Trans. Inf. Forensics Secur. 15 (2020) 3540–3552.

[31] T. Hospedales, A. Antoniou, P. Micaelli, A. Storkey, Meta-Learning in Neural Networks: A Survey, IEEE Trans. Pattern Anal. Mach. Intell. 44 (9) (2022) 5149–5169.
[32] E. Lee, W. Rhee, Individualized short-term electric load forecasting with deep neural network based transfer learning and meta learning, IEEE Access 9 (2021) 15413–15425.
[33] X. Bing, Critical infrastructure protection based on memory-augmented meta-learning framework, Neural Comput. Appl. 32 (23) (2020) 17197–17208.
[34] C. Lemke, M. Budka, B. Gabrys, Metalearning: a survey of trends and technologies, Artif. Intell. Rev. 44 (1) (2015) 117–130.
[35] C. Giraud-Carrier, Metalearning-a tutorial, in: Tutorial at the 7th International Conference on Machine Learning and Applications (ICMLA), San Diego, California, USA, 2008.
[36] B. Lake, R. Salakhutdinov, J. Gross, J. Tenenbaum, One shot learning of simple visual concepts, in: Proceedings of the Annual Meeting of the Cognitive, Science Society vol. 33, 2011, pp. 2568–2573.
[37] L. Fei-Fei, R. Fergus, P. Perona, One-shot learning of object categories, IEEE Trans. Pattern Anal. Mach. Intell. 28 (4) (2006) 594–611.
[38] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (10) (2009) 1345–1359.
[39] N. Abbas, Y. Zhang, A. Taherkordi, T. Skeie, Mobile edge computing: a survey, IEEE Internet Things J. 5 (1) (2017) 450–465.
[40] S. Safavat, N.N. Sapavath, D.B. Rawat, Recent advances in mobile edge computing and content caching, Digit. Commun. Netw. 6 (2) (2020) 189–194.
[41] K. Sha, T.A. Yang, W. Wei, S. Davari, A survey of edge computing-based designs for iot security, Digit. Commun. Netw. 6 (2) (2020) 195–202.
[42] R.-F. Liao, H. Wen, J. Wu, F. Pan, A. Xu, H. Song, F. Xie, Y. Jiang, M. Cao, Security enhancement for mobile edge computing through physical layer authentication, IEEE Access 7 (2019) 116390–116401.
[43] R. Atat, L. Liu, H. Chen, J. Wu, H. Li, Y. Yi, Enabling cyber-physical communication in 5g cellular networks: challenges, spatial spectrum sensing, and cyber-security, IET Cyber-Phys. Syst.: Theor. Appl. 2 (1) (2017) 49–54.
[44] L. Huang, L. Zhang, S. Yang, L.P. Qian, Y. Wu, Meta-learning based dynamic computation task offloading for mobile edge computing networks, IEEE Commun. Lett. 25 (5) (2020) 1568–1572.
[45] M. Min, L. Xiao, Y. Chen, P. Cheng, D. Wu, W. Zhuang, Learning-based computation offloading for iot devices with energy harvesting, IEEE Trans. Veh. Technol. 68 (2) (2019) 1930–1941.
[46] L. Huang, S. Bi, Y.-J.A. Zhang, Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks, IEEE Trans. Mobile Comput. 19 (11) (2019) 2581–2593.
[47] X. Wang, Y. Han, V.C. Leung, D. Niyato, X. Yan, X. Chen, Convergence of edge computing and deep learning: a comprehensive survey, IEEE Commun. Surv. Tutor. 22 (2) (2020) 869–904.
[48] A. Ghasempour, J. Lou, Advanced metering infrastructure in smart grid: requirements, challenges, crchitectures, technologies, and optimizations, in: Smart Grids: Emerging Technologies, Challenges and Future Directions, Nova Science Publishers, 2017, pp. 1–8.
[49] W. Lei, H. Wen, J. Wu, W. Hou, Maddpg-based security situational awareness for smart grid with intelligent edge, Appl. Sci. 11 (7) (2021) 3101.
[50] J. Li, M. Hu, Continuous model adaptation using online meta-learning for smart grid application, IEEE Transact. Neural Networks Learn. Syst. 32 (8) (2021) 3633–3642.
[51] S. Ding, C. Dong, T. Zhao, L. Koh, X. Bai, J. Luo, A meta-learning based multimodal neural network for multistep ahead battery thermal runaway forecasting, IEEE Trans. Ind. Inf. 17 (7) (2021) 4503–4511.
[52] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Y. Bengio, Learning phrase representations using rnn encoder-decoder for statistical machine translation, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, 2014, pp. 1724–1734.
[53] M. Schuster, K.K. Paliwal, Bidirectional recurrent neural networks, IEEE Trans. Signal Process. 45 (11) (1997) 2673–2681.
[54] S. Sharma, R. Gupta, Intrusion detection system: a review, Int. J. Secur. Appl. 9 (5) (2015) 69–76.
[55] L. Bilge, T. Dumitraş, Before we knew it: an empirical study of zero-day attacks in the real world, in: Proceedings of the 2012 ACM Conference on Computer and Communications Security, ACM, 2012, pp. 833–844.
[56] M. Chalé, N.D. Bastian, J. Weir, Algorithm selection framework for cyber attack detection, in: Proceedings of the 2nd ACM Workshop on Wireless Security and Machine Learning, 2020, pp. 37–42.
[57] M. Ahsan, R. Gomes, M. Chowdhury, K.E. Nygard, et al., Enhancing machine learning prediction in cybersecurity using dynamic feature selector, J. Cybersecur. Priv. 1 (1) (2021) 199–218.
[58] S.R. Safavian, D. Landgrebe, A survey of decision tree classifier methodology, IEEE Trans. Syst. Man, Cybern. 21 (3) (1991) 660–674.
[59] S. Chen, H. Wen, J. Wu, W. Lei, W. Hou, W. Liu, A. Xu, Y. Jiang, Internet of things based smart grids supported by intelligent edge computing, IEEE Access 7 (2019) 74089–74102.
[60] N.H.N. Zulkipli, A. Alenezi, G.B. Wills, Iot forensic: bridging the challenges in digital forensic and the internet of things, in: International Conference on Internet of Things, Big Data and Security vol. 2, SCITEPRESS, 2017, pp. 315–324.
[61] J. Pan, Iot network behavioral fingerprint inference with limited network traces for cyber investigation, in: IEEE International Conference on Artificial Intelligence in Information and Communication (ICAIIC), IEEE, 2021, pp. 263–268.
[62] S. Xingjian, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, W.-c. Woo, Convolutional lstm network: a machine learning approach for precipitation nowcasting, in: Advances in Neural Information Processing Systems, 2015, pp. 802–810.
[63] S.-C. Lin, C.-J. Chen, T.-J. Lee, A multi-label classification with hybrid label-based meta-learning method in internet of things, IEEE Access 8 (2020) 42261–42269.
[64] D. Mishra, B. Naik, P.B. Dash, J. Nayak, Sem: stacking ensemble meta-learning for iot security framework, Arabian J. Sci. Eng. 46 (4) (2021) 3531–3548.
[65] M. Khonji, Y. Iraqi, A. Jones, Phishing detection: a literature survey, IEEE Commun. Surv. Tutor. 15 (4) (2013) 2091–2121.
[66] H. Zhu, Online meta-learning firewall to prevent phishing attacks, Neural Comput. Appl. 32 (2020) 17137–17147.
[67] M. Amiri, L. Mohammad-Khanli, Survey on prediction models of applications for resources provisioning in cloud, J. Netw. Comput. Appl. 82 (2017) 93–113.
[68] M. Kim, Ml/cgan: network attack analysis using cgan as meta-learning, IEEE Commun. Lett. 25 (2) (2021) 499–502.
[69] Y. Pan, P. Zhou, Y. Yan, A. Agrawal, Y. Wang, D. Guo, S. Goel, New insights into the methods for predicting ground surface roughness in the age of digitalisation, Precis. Eng. 67 (2021) 393–418.
[70] J.K.-F. Pun, Improving Credit Card Fraud Detection Using a Meta-Learning Strategy, Ph.D. thesis, University of Toronto, 2011.
[71] T.A. Olowookere, O.S. Adewale, A framework for detecting credit card fraud with cost-sensitive meta-learning ensemble approach, Sci. Afr. 8 (2020), e00464.
[72] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: a survey, ACM Comput. Surv. 41 (3) (2009) 1–58.
[73] M.M. Breunig, H.-P. Kriegel, R.T. Ng, J. Sander, Lof: identifying density-based local outliers, in: Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, 2000, pp. 93–104.
[74] F.T. Liu, K.M. Ting, Z.-H. Zhou, Isolation forest, in: IEEE International Conference on Data Mining (ICDM), IEEE, 2008, pp. 413–422.
[75] T.K. Tran, H. Sato, M. Kubo, One-shot learning approach for unknown malware classification, in: Proceedings of the 2018 5th Asian Conference on Defense Technology (ACDT), IEEE, 2018, pp. 8–13.
[76] M.C. Kang, H.K. Kim, Rare malware classification using memory augmented neural networks, J. Korea Inst. Inform. Secur. Cryptol. 28 (4) (2018) 847–857.
[77] T.K. Tran, H. Sato, M. Kubo, Image-based unknown malware classification with few-shot learning models, in: Proceedings of the 2019 7th International Symposium on Computing and Networking Workshops (CANDARW), IEEE, 2019, pp. 401–407.
[78] Z. Tang, P. Wang, J. Wang, Convprotonet: deep prototype induction towards better class representation for few-shot malware classification, Appl. Sci. 10 (8) (2020) 2847.
[79] J. Shin, Cross-domain meta-learning for bug finding in the source codes with a small dataset, in: Proceedings of the European Interdisciplinary Cybersecurity Conference, 2020, pp. 1–6.
[80] B. McMahan, E. Moore, D. Ramage, S. Hampson, B.A. y Arcas, Communication-efficient learning of deep networks from decentralized data, in: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics vol. 54, PMLR, 2017, pp. 1273–1282.
[81] T. Li, A.K. Sahu, A. Talwalkar, V. Smith, Federated learning: challenges, methods, and future directions, IEEE Signal Process. Mag. 37 (3) (2020) 50–60.
[82] Z. Chen, W. Liao, K. Hua, C. Lu, W. Yu, Towards asynchronous federated learning for heterogeneous edge-powered internet of things, Digit. Commun. Netw. 7 (3) (2021) 317–326.
[83] F. Chen, M. Luo, Z. Dong, Z. Li, X. He, Federated Meta-Learning with Fast Convergence and Efficient Communication, 2019 arXiv:1802.07876.
[84] S. Lin, G. Yang, J. Zhang, A collaborative learning framework via federated meta-learning, in: IEEE 40th International Conference on Distributed Computing Systems (ICDCS), IEEE, 2020, pp. 289–299.
[85] A.-m. Yang, Y.-x. Zhuansun, Prediction of compressive strength based on visualization of pellet microstructure data, J. Iron Steel Res. Int. 28 (6) (2021) 651–660.
[86] S. Mei, X. Zhu, Using machine teaching to identify optimal training-set attacks on machine learners, in: Proceedings of the AAAI Conference on Artificial Intelligence vol. 29, 2015, pp. 2871–2877.
[87] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, R. Fergus, Intriguing properties of neural networks (2013) arXiv:1312.6199.
[88] L. Muñoz-González, B. Biggio, A. Demontis, A. Paudice, V. Wongrassamee, E.C. Lupu, F. Roli, Towards poisoning of deep learning algorithms with back-gradient optimization, in: Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security, 2017, pp. 27–38.
[89] D. Zügner, A. Akbarnejad, S. Günnemann, Adversarial attacks on neural networks for graph data, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018, pp. 2847–2856.
[90] D. Zügner, S. Günnemann, Adversarial attacks on graph neural networks via meta learning, in: International Conference on Learning Representations, ICLR), 2018.
[91] Y. Zhou, J. Ren, D. Dou, R. Jin, J. Zheng, K. Lee, Robust meta network embedding against adversarial attacks, in: IEEE International Conference on Data Mining (ICDM), IEEE, 2020, pp. 1448–1453.
[92] R. Edmunds, N. Golmant, V. Ramasesh, P. Kuznetsov, P. Patil, R. Puri, Transferability of adversarial attacks in model-agnostic meta-learning, in: Deep Learning and Security Workshop (DLSW) in Singapore, 2017.
[93] I.-T. R. ITU-T Y, 3001, Future Networks: Objectives and Design Goals, ITU, Geneva, Switzerland, 2011.
[94] W. Bhebe, O.P. Kogeda, Shilling attack detection in collaborative recommender systems using a meta learning strategy, in: IEEE International Conference on

Emerging Trends in Networks and Computer Communications (ETNCC), IEEE, 2015, pp. 56–61.

[95] M. Dibaei, X. Zheng, K. Jiang, R. Abbas, S. Liu, Y. Zhang, Y. Xiang, S. Yu, Attacks and defences on intelligent connected vehicles: a survey, Digit. Commun. Netw. 6 (4) (2020) 399–421.

[96] Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, Timothy Lillicrap, in: Proceedings of The 33rd International Conference on Machine Learning, 48, PMLR, 2016, pp. 1842–1850.