symmetry

MDPI

Article

Artificial Intelligence in the Cyber Domain: Offense and Defense

Thanh Cong Truong [1,2,*,†], Quoc Bao Diep [1,†] and Ivan Zelinka [1,3,†]

[1] Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, 17. listopadu 2172/15, Ostrava-Poruba, 708 00 Ostrava, Czech Republic; quoc.bao.diep.st@vsb.cz (Q.B.D.); ivan.zelinka@vsb.cz or ivan.zelinka@tdt.edu.vn (I.Z.)

[2] Faculty of Information Technology, University of Finance-Marketing, Ho Chi Minh City, Vietnam

[3] Modeling Evolutionary Algorithms Simulation and Artificial Intelligence, Faculty of Electrical & Electronics Engineering, Ton Duc Thang University, Ho Chi Minh City, Vietnam

* Correspondence: cong.thanh.truong.st@vsb.cz or ttcong@ufm.edu.vn; Tel.: +420-774-820-835

† These authors contributed equally to this work.

Received: 30 December 2019; Accepted: 8 February 2020; Published: 4 March 2020

Abstract: Artificial intelligence techniques have grown rapidly in recent years, and their applications in practice can be seen in many fields, ranging from facial recognition to image analysis. In the cybersecurity domain, AI-based techniques can provide better cyber defense tools and help adversaries improve methods of attack. However, malicious actors are aware of the new prospects too and will probably attempt to use them for nefarious purposes. This survey paper aims at providing an overview of how artificial intelligence can be used in the context of cybersecurity in both offense and defense.

Keywords: cybersecurity; artificial intelligence; machine learning; deep learning; bio-inspired computing; systems security

1. Introduction

Cybersecurity involves the devising of defense strategies that preserve computing resources, networks, programs, and data from unauthorized access, change, or destruction. Due to the dramatic advances in information and communication technologies, new cybersecurity threats are emerging and changing rapidly. Cybercriminals are adopting new and sophisticated techniques that increase the speed and scale of their attacks. Hence, there is a requirement for more flexible, adaptable, and robust cyber defense systems that are capable of detecting a wide variety of threats in real-time. In recent years, the adoption of artificial intelligence (AI) techniques has been rising and maintaining a crucial role in cyber threat detection and prevention.

While the concept of AI was proposed in the 1950s, in recent years, it has grown at a significant pace and is now influencing all aspects of communities and occupations. Many areas benefit from AI, such as gaming, natural language processing, health care, manufacturing, education, and others. This trend is also affecting the cybersecurity field where AI has been utilized for both attacking and defending in the cyberspace. On the offense side, cyber threats can employ AI to improve the sophistication and scope of their attacks. On the defense side, AI is utilized to enhance the defense strategies, so that the defense systems become more robust, flexible, and efficient, which involves being adaptive with changes in the environment to decrease the impacts occurred.

Recently, researchers presented several surveys in the domain of AI and cybersecurity. However, some of them just focused on the adopting machine learning methods for cyber problems such as those in [1–4]. Other research [5,6] just focused on deep learning methods. Additionally, there is a lack of literature dealing with the nefarious use of AI.

Apruzzese et al. [7] performed a survey on ML and DL methods for cyber security. Nevertheless, their research was just covering attacks related particularly to network intrusion detection, malware investigation, and spam identification.

The author in [8] discussed the intersection of AI and cybersecurity. More particularly, the paper reviewed some ML and DL approaches to counter against cyber attacks. What is more, the author introduced the possibility of attacking the AI model. Nevertheless, the paper just discussed adversarial attacks and ignored other kinds of attack using the AI model, such as poisoning data, and the extraction model.

Another approach by the authors in [4] pointed out the differences between traditional ML and DL methods for cybersecurity. However, their survey just concentrated on intrusion detection.

Based on the above circumstances, this survey paper pursues a two-fold goal. The first is to carry out an exploration of the impact of AI in cybersecurity. The second is to replenish the literature with recent reviews on cyber applications of AI methods.

The main contributions of this survey are listed as follows:

- To present the impact of AI techniques on cybersecurity: we provide a brief overview of AI and discuss the impact of AI in the cyber domain.
- Applications of AI for cybersecurity: we conduct a survey of the applications of AI for cybersecurity, which covers a wide-range cyber attack types.
- Discussion on the potential security threats from adversarial uses of AI technologies: we investigate various potential threats and attacks may arise through the use of AI systems.
- Challenges and future directions: We discuss the potential research challenges and open research directions of AI in cybersecurity.

The remainder of this paper proceeds as follows. Section 2 describes the research methodology. Section 3 presents a brief overview of AI and discusses the role of AI in cybersecurity. Section 4 gives a brief overview of AI methodology for cybersecurity. Section 5 focuses on cyber applications of AI methods. Section 6 discusses various potential threats that adopt AI techniques. Challenges and open research directions are discussed in Section 7. Section 8 provides short discussions about the role of AI, and compares our work with existing surveys. Section 9 concludes the paper.

2. Research Methodology

To get a comprehensive overview of the junction between AI and cyber security, we used four databases: Web of Science, Scopus, IEEE Xplore, and ACM digital library. Alongside that, the Google Scholar search engine was also utilized. A set of keywords related to the topic have been used in these databases. To enhance the search results, the authors refined different keywords and keyword mixtures for each search engine to obtain the highest coverage.

In a second step, we used a filter based on the obtained results. The search results were limited only to the existing papers published in the last four years, because the purpose of this paper is to discover the most recent research trends of AI in cybersecurity. Next, the results were sorted by the number of citations, and manuscripts that had more than five citations were selected. On the other hand, recently published papers which had less than five citations but had novel approaches were also chosen. After that, the materials which met the following criteria were excluded:

- Papers which had titles belonging to subjects outside the scope of this research.
- Books, patent documents, technical reports, citations.
- Papers which were not written in English.

In the third step, we examined the abstracts and the conclusions for relevant data. Through this step, the authors confirmed whether the classified papers matched the main topic of the junction between AI and cybersecurity. Consequently, those papers which were the most relevant to the task were chosen.

3. The impact of AI on Cybersecurity

Defining AI can take two approaches. First, it is a science that strives to discover the nature of intelligence and develop smart machines in which scientists apply information, logic, self-learning, and determination to make machines become intelligent. To put it simply, humans create machines with intelligence. This intelligence can think, learn, decide, and work while trying to solve a problem, as a human intellect does. On the other hand, scientists define AI as a science that researches and develops methods for resolving complexity problems that are impossible to be resolved without adopting intelligence. For example, scientists can build an AI system for real-time analysis and decision making based on enormous amounts of data. In recent years, AI has resulted in advances in many scientific and technological fields, such as computerized robots, image recognition, natural language processes, expert systems, and others.

The rapid development of computing technology and the internet has a significant impact on people's daily lives and work. Unfortunately, it also caused many new cybersecurity challenging issues: First, the explosion of data makes manual analysis impractical. Second, threats are growing at a high rate, which also means that new, short-lived species and highly adaptive threats become commonplace. Third, at present, the threats compromise various techniques for propagation, infection, and evasion; therefore, they are hard to detect and predict. Moreover, the expense to prevent threats also should be considered. It takes a lot of time, money, and effort to generate and implement an algorithm. Additionally, employing or training specialists in the field is hard and expensive. What is more, many threat variations emerge and spread continuously. Hence, AI-based methods are expected to cope with these cybersecurity issues.

3.1. The Positive Uses of AI

In the field of cybersecurity, AI is already being used to advance defensive capabilities. Based on its powerful automation and data analysis capabilities, AI can be used to analyze large amounts of data with efficiency, accuracy, and speed. An AI system can take advantage of what it knows and understand the past threats to identify similar attacks in the future, even if their patterns change. Undoubtedly, artificial intelligence has several advantages when it comes to cybersecurity in the following aspects:

- AI can discover new and sophisticated changes in attack flexibility: Conventional technology is focused on the past and relies heavily on known attackers and attacks, leaving room for blind spots when detecting unusual events in new attacks. The limitations of old defense technology are now being addressed through intelligent technology. For example, privileged activity in an intranet can be monitored, and any significant mutation in privileged access operations can denote a potential internal threat. If the detection is successful, the machine will reinforce the validity of the actions and become more sensitive to detecting similar patterns in the future. With a larger amount of data and more examples, the machine can learn and adapt better to detect anomalous, faster, and more accurate operations. This is especially useful while cyber-attacks are becoming more sophisticated, and hackers are making new and innovative approaches.
- AI can handle the volume of data: AI can enhance network security by developing autonomous security systems to detect attacks and respond to breaches. The volume of security alerts that appear daily can be very overwhelming for security groups. Automatically detecting and responding to threats has helped to reduce the work of network security experts and can assist in detecting threats more effectively than other methods. When a large amount of security data is created and transmitted over the network every day, network security experts will gradually have difficulty tracking and identifying attack factors quickly and reliably. This is where AI can help, by expanding the monitoring and detection of suspicious activities. This can help network security personnel react to situations that they have not encountered before, replacing the time-consuming analysis of people.

- An AI security system can learn over time to respond better to threats: AI helps detect threats based on application behavior and a whole network's activity. Over time, AI security system learns about the regular network of traffic and behavior, and makes a baseline of what is normal. From there, any deviations from the norm can be spotted to detect attacks.

AI techniques seem an up-and-coming area of research that enhances the security measures for cyberspace. Many AI methods are being used to deal with threats, including computational intelligence, neural networks, intelligent agents, artificial immune systems, data mining, pattern recognition, heuristics, ML, DL, and others. However, among these techniques, ML and DL attracted a lot of attention recently and obtained the most achievements in combating against cyber-threats.

3.2. Drawbacks and Limitations of Using AI

The advantages highlighted above are just a fraction of the potential of how AI can assist cybersecurity, but the application of this technology has some limitations, as described below.

- Data sets: Creating an AI system demands a considerable number of input samples, and obtaining and processing the samples can take a long time and a lot of resources.
- Resource requirements: Building and maintaining the fundamental system needs an immense amount of resources, including memory, data, and computing power. What is more, skilled resources necessary to implement this technology require a significant cost.
- False alarms: Frequent false alarms are an issue for end-users, disrupting business by potentially delaying any necessary response and generally affecting efficiency. The process of fine-tuning is a trade-off between reducing false alarms and maintaining the security level.
- Attacks on the AI-based system: Attackers can use various attack techniques that target AI systems, such as adversarial inputs, data poisoning, and model stealing.

One important aspect to be taken into account is the nefarious use of AI. This technology will also be used as a way to improve threats. For example, malicious actors can leverage the ML technique to generate a hard-to-detect malware variant with machine speed. What is more, AI might be able to personalize the phishing scheme better and raise the scale of the attack, making the attack more likely to succeed. More detail about this matter discussed in Section 6.

4. AI Methodology for Cybersecurity

In this section, the authors give an overview of the learning algorithms, an essential concept of AI. Furthermore, we present a brief introduction about ML, DL, and bio-inspired computation methods that are frequently utilized in the area of cybersecurity.

4.1. Learning Algorithms

AI is a branch of computer science that seeks to produce a new type of intelligent automaton that responds like human intelligence. To achieve this goal, machines need to learn. To be more precise, we need to train the computer by using the learning algorithms. Generally, learning algorithms help to enhance performance in accomplishing a task through learning and training from experience. There are currently three major types of learning algorithms which we use to train machines:

- Supervised learning: This type requires a training process with a large and representative set of data that has been previously labeled. These learning algorithms are frequently used as a classification mechanism or a regression mechanism.
- Unsupervised learning: In contrast to supervised learning, unsupervised learning algorithms use unlabeled training datasets. These approaches are often used to cluster data, reduce dimensionality, or estimate density.
- Reinforcement learning: Reinforcement learning is a type of learning algorithm that learns the best actions based on rewards or punishment. Reinforcement learning is useful for situations where data is limited or not given.

4.2. Machine Learning Methods

Machine learning (ML) is a branch of AI that aims to empower systems by utilizing data to learn and improve without being explicitly programmed. ML has strong ties to mathematical techniques that enable a process of extracting information, discovering patterns, and drawing conclusions from data. There are different types of the ML algorithm, but they can generally be classified into three main categories: supervised learning, unsupervised learning, and reinforcement learning. In the computer security domain, the standard ML algorithms are decision trees (DT), support vector machines (SVM), Bayesian algorithms, k-nearest neighbor (KNN), random forest (RF), association rule (AR) algorithms, ensemble learning (EL), k-means clustering, and principal component analysis (PCA).

4.3. Deep Learning Methods

Deep learning (DL) is a sub-field of ML, and it uses data to teach computers how to do things only humans are capable of at that time. Its motivation lies in the working mechanisms of the human brain and neurons for processing signals. The core of deep learning is that if we construct more extensive neural networks and train them with as much data as possible, their performance continues to increase. The most important advantage of DL over the conventional ML is its superior performance in large datasets. Similarly to ML methods, DL methods also have supervised learning, unsupervised learning, and reinforcement learning. The benefit of DL is the leverage of unsupervised learning to select feature automatically. The typical DL algorithms frequently utilized in the cybersecurity domain are: feed forward neural networks (FNN), convolutional neural networks (CNNs), recurrent neural networks (RNN), deep belief networks (DBNs), stacked autoencoders (SAE), generative adversarial networks (GANs), restricted Boltzmann machines (RBMs), and ensemble of DL networks (EDLNs).

4.4. Bio-Inspired Computation Methods

Bio-inspired computation is a branch of AI which emerged as one of the most studied during recent years. It is a collection of intelligent algorithms and methods that adopt bio-inspired behaviors and characteristics to solve a wide range of complex academic and real domain problems. Among many biological-inspired methods, the following techniques are most commonly used in the cybersecurity domain: genetic algorithms (GA), evolution strategies (ES), ant colony optimization (ACO), particle swarm optimization (PSO), and artificial immune systems (AIS).

5. AI-Based Approaches for Defending Against Cyberspace Attacks

Recently, scientists proposed numerous techniques that have utilized AI methods to detect or categorize malware, detect network intrusions, phishing, and spam attacks; counter Advanced persistent threat (APT); and identify domain generated by doamain generation algorithms (DGAs). In this section, we category these literature into four main groups: malware identification; network intrusion detection; phishing and SPAM identification; and other, which compromises countering APT and identifying DGAs. Figure 1 illustrates the primary areas of utilizing AI for cybersecurity.

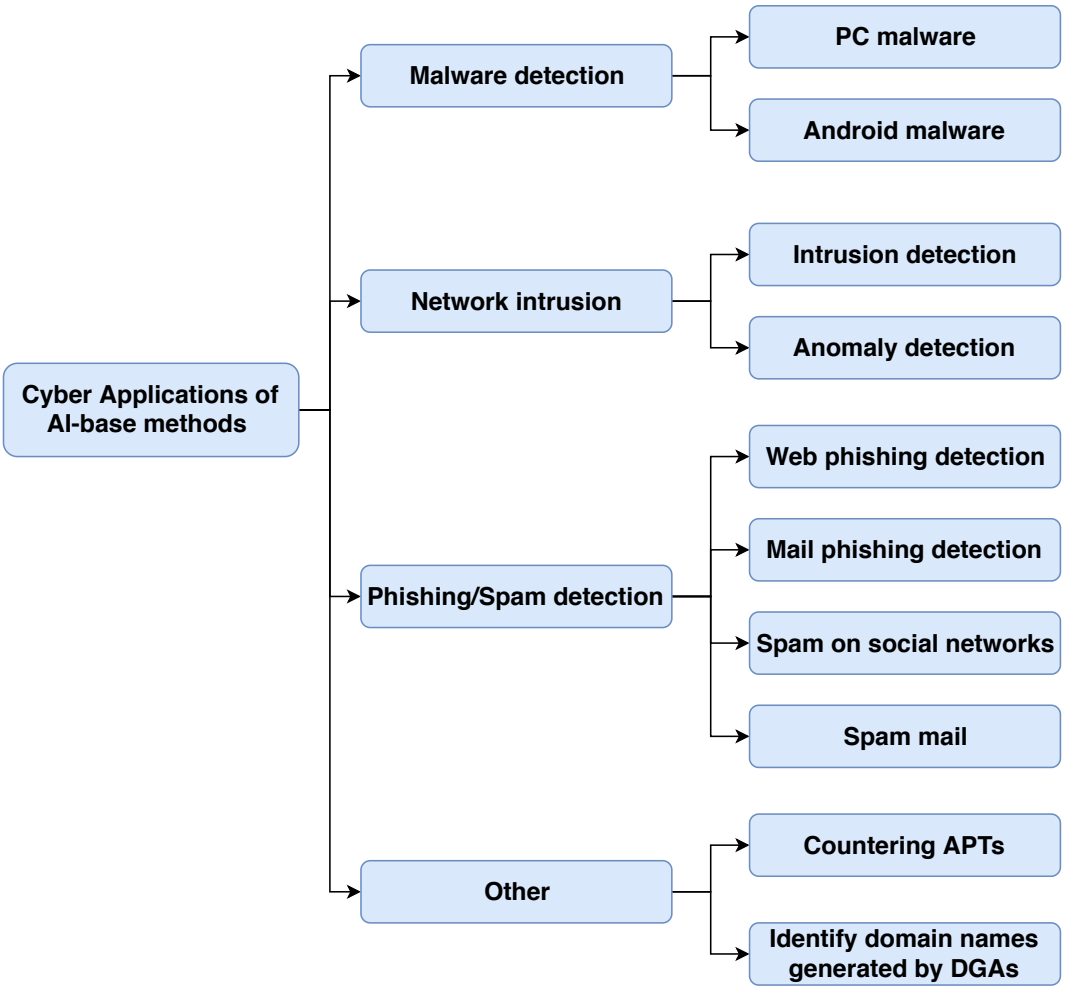

Figure 1. Main branches of cybersecurity applications adopting AI techniques.

5.1. Malware Identification

Malware is a general term for many types of malicious software, such as viruses, worms, trojan horses, exploits, botnet, retroviruses, and today, malware is a popular method of cyber-attack. Malware's impact on digital society is enormous, so a considerable amount of research about adopting AI techniques has been done to prevent and mitigate malware. The most recent and noteworthy contributions utilize intelligence for malware detection and prevention—described as follows.

In [9], the authors adopted ML to create an online framework for hardware-assisted malware detection based on virtual memory access patterns. The proposed method used logistic regression, a support vector machine, and a random forest classifier, and performed on the RIPE benchmark suite for the experiments. The authors reported that the framework has a true positive rate of 99% with a less than 5% false positive rate. Meanwhile, the scholars in [10] presented a framework for classifying and detecting malicious software using data mining and ML classification. In that work, both signature-based and anomaly-based features were analyzed for detection. Experimental results showed that the proposed method outperformed other similar methods.

Another approach [11] used operational codes (OpCode), k-nearest neighbors (KNN), and a support vector machine (SVM) as ML classifiers to classify malware. The OpCode was represented as a graph and embedded into eigenspace; then, one classifier or an ensemble of classifiers were utilized to classify each vector as malware or benign. The empirical result showed that the proposed model is efficient with a low false alarm rate and high detection rate.

Later, Ye et al. [12] built a deep learning architecture for intelligent malware detection. In this work, they utilized an AutoEncoder stacked up with multilayer restricted Boltzmann machines (RBMs)

to detect unknown malware. The author claimed that heterogeneous deep learning framework could improve the overall performance in malware detection compared with traditional shallow learning methods and deep learning methods.

A recent trend of research in malware detection focused on mobile malware in general and Android malware in particular. Machine learning, along with deep learning, was a significant breakthrough in this area. In [13], a deep convolutional neural network (CNN) was adopted to identify malware. The raw opcode sequence from a disassembled program was used to classify malware. The authors in [14] utilized a support vector machine (SVM) and the most significant permissions from all of the permission data to distinguish between benign and malicious apps. In [15], the authors presented novel ML algorithms, namely, rotation forest, for malware identity. An artificial neural network (ANN) and the raw sequences of API method calls were utilized in [16] to detect Android malware. A recent study by Wang et al. [17] introduced a hybrid model based on deep autoencoder (DAE) and a convolutional neural network (CNN) to raise the accuracy and efficiency of large-scale Android malware detection.

Another research direction that attracted the attention of scientists was the use of bio-inspired methods for malware classification. These techniques were mainly used for feature optimization and optimizing the parameter for the classifiers. For example, particle swarm optimization (PSO) was adopted in [18–20]; the genetic algorithm (GA) was utilized in [21,22] to enhance the effectiveness of a malware detection system.

Table 1 abstracts some characteristics of the discussed malware identification approaches, concerning the focus areas, techniques, features, datasets, and validation metrics used to evaluate the models' performances. For the validation metrics, we present the best performing method in the paper. For the multi-task model, we present the evaluation measures of all tasks, if they exist, given by the "/" symbol. The acronyms for this table are described in Table 2.

Table 1. Selected literature of AI-based approaches in malware investigation.

References	Year	Focus	Tech.	Features	Dataset	Validation Metrics
[9]	2017	PC malware	SVM,RF Logistic regression	MAP's feature sets	RIPE	DR: 99% FPR: 5%
[10]	2017	PC malware	BAM, MLP	N-gram, Windows API calls	Self collection: 52,185 samples	ACC: 98.6% FPR: 2%
[11]	2017	PC malware	KNN, SVM	OpCode graph	Self collection: 22,200 samples	ACC, FPR
[13]	2017	Android malware	CNN	Opcode sequence	GNOME, McAfee Labs	ACC: 98%/80%/87%, F-score: 97%/78%/86%
[19]	2017	Android malware	ANF, PSO	Permissions, API Calls	Self collection: 500 samples	ACC: 89%
[21]	2017	Botnet	C4.5, GA	Multi features	ISOT , ISCX	DR: 99.46%/95.58% FPR: 0.57%/ 2.24%
[12]	2018	PC malware	AutoEncoder, RBM	Windows API calls	Self collection: 20,000 samples	ACC: 98.2%
[14]	2018	Android malware	SVM, DT	Significant permissions	Self collection: 54,694 samples	ACC: 93.67% FPR: 4.85%
[15]	2018	Android malware	Rotation Forest	Permissions, APIs, system events	Self collection: 2,030 smaples	ACC: 88.26%
[16]	2018	Android malware	ANN	API call	Malgenome, Drebin, Maldozer	F1-Score: 96.33% FPR: 3.19%
[18]	2018	Android malware	PSO, RF, J48, KNN, MLP, AdaBoost	Permissions	Self collection: 8500 samples	TPR: 95.6% FPR: 0.32%

Table 1. *Cont.*

References	Year	Focus	Tech.	Features	Dataset	Validation Metrics
[17]	2019	Android malware	DAE, CNN	Permissions, filtered intents, API calls, hardware features, code related patterns	Self collection: 23000 samples	ACC: 98.5%/98.6% FPR: 1.67%/1.82%
[20]	2019	Android malware	PSO, Bayesnet, Naïve Bayes, SMO, DT, RT, RF J48, MLP	Permissions	UCI, KEEL, Contagiodump, Wang's repository	ACC: 79.4%/47.6%/ 82.9%/94.1%/ 100%/77.9%
[22]	2019	Android malware	SVM, ANN	App Components, Permissions	Self collection: 44,000 samples	ACC: 95.2%/96.6%

Table 2. The acronyms used in Table 1.

ACC: Accuracy	CNN: Convolutional neural network
FPR: False positive rate	ANF: Adaptive neural fuzzy
DR: Detection rate	GA: Genetic Algorithm
RF: Random forest	RBMs: Restricted Boltzmann machines
SVM: Support vector machine	DT: Decision tree
MLP: multilayer perceptron	GP: Genetic programming
BAM: binary associative memory	DT: decision tree
KNN: k-nearest neighbors	DAE: Deep auto-encoder

5.2. Intrusion Detection

An intrusion detection system (IDS) is a system that is supposed to protect the system from possible incidents, violations, or imminent threats. AI-based techniques are appropriate for developing IDS, and outperform other techniques because of their flexibility, adaptability, rapid calculations, and quick learning. Hence, many researchers studied intelligent methods to improve the performance of IDS. The focus was on developing optimized features and improving the classifiers to reduce the false alarms. Some recent notable studies are listed as follows.

Al-Yaseen et al. [23] combined a support vector machine (SVM) and an extreme learning machine with modified k-means as a model for IDS. Using the KDD'99 Cup dataset, their model archived a result of up to 95.75% accuracy and 1.87% false alarms. Meanwhile, Kabir et al. [24] introduced a method for an intrusion detection system based on sampling with a least square support vector machine (LS-SVM). The proposed methodology was validated through the KDD'99 Cup dataset and obtained a realistic performance in terms of accuracy and efficiency.

The authors in [25] introduced a fuzziness based semi-supervised learning approach for IDS. In their work, they utilized unlabeled samples assisted with a supervised learning algorithm to enhance the performance of the classifier. The algorithm was tested on the KDD'99 Cup dataset and outperformed other comparative algorithms.

Later, Shone et al. [26] proposed a novel deep learning-based intrusion detection method called nonsymmetric deep autoencoder (NDAE). The authors used TensorFlow and evaluated their method by using KDD Cup '99 and NSL-KDD datasets. They have claimed that their model achieved an accuracy of 97.85%.

Another approach using genetic algorithms (GA) and fuzzy logic for network intrusion detection is presented by Hamamoto et al. [27]. The GA is used to create a digital signature of a network segment using glow analysis (DSNSF), a prediction of the network's traffic behavior for a given time interval. Additionally, the fuzzy logic approach is adopted to assess whether an instance represents an anomaly or not. The evaluation was conducted by using real network traffic from a university and obtained an accuracy of 96.53% and a false alarm of 0.56%.

One point to be taken into account is that the use of swarm intelligence (SI) for IDS. Botes et al. [28] presented a new method, namely, ant tree miner (ATM) classification, which is a decision tree

using ACO instead of conventional techniques, such as C4.5 and CART [29], for intrusion detection. Using NSL-KDD datasets, their approach achieved the accuracy of 65% and a false alarm rate pf 0%.

In a later study [30], the authors presented an IDS using binary PSO and KNN. The proposed method consists of feature selection and classification steps. Based on the results obtained, the algorithm showed excellent performance, and the proposed hybrid algorithm raised the accuracy generated by KNN by up to 2%. Meanwhile, Ali et al. [31] introduced a learning model for a fast learning network (FLN) based on PSO named PSO-FLN, and then the model was utilized for the problem of IDS. The PSO-FLN model was tested on the KDD'99 Cup datasets and achieved the highest testing accuracy compared to other meta-heuristic algorithms.

In the recent study by Chen et al. [32], a multi-level adaptive coupled intrusion detection method combining white list technology and machine learning was presented. The white list was used to filter the communication, and the machine learning model was used to identify abnormal communication. In this article, the adaptive PSO algorithm and the artificial fish swarm (AFS) algorithm were used to optimize the parameters for the machine learning model. The method was tested on KDD'99 Cup, Gas Pipeline, and industrial field datasets. The empirical result showed that the proposed model is efficient with various attack types.

In [33], the authors introduced the Fuzzified Cuckoo based clustering technique for anomaly detection. The technique consists of two phases: the training phase and the detection phase. In the training phase, cuckoo search optimization (CSO), k-means clustering, and decision tree criterion (DTC) were combined to evaluate the distance functions. In the detection phase, a fuzzy decisive approach was utilized to identify the anomalies based on input data and previously computed distance functions. Experimental results showed that the model was effective with an accuracy rate of 97.77% and a false alarm rate of 1.297%.

Meanwhile, the authors in [34] incorporated artificial bee colony and artificial fish swarm algorithms to cope with the complex IDS problems. In this work, a hybrid classification method based on the ABC and AFS algorithms was proposed to improve the detection accuracy of IDS. The NSL-KDD and UNSW-NB15 datasets were used to evaluate the performance of the method. Based on the results obtained, the proposed model was efficient with a low false alarm rate and high accuracy rate.

In later research, Garg et al. [35] proposed a hybrid model for network anomaly detection in cloud environments. The model utilized gray wolf optimization (GWO) and a convolutional neural network (CNN) for feature extraction and identifying the anomalies in real-time network traffic streams. The empirical result showed that the proposed model was efficient with a low false alarm rate and high detection rate.

Another approach [36] presented a hybrid IDS utilizing spark ML and the convolutional-LSTM network. The ISCX-UNB dataset was used to evaluate the performance of the method. Based on the results obtained, the proposed model obtained a significant result and outperformed the compared method.

In reference [37], the authors adopted the firefly algorithm for feature selection and the C4.5 Bayesian networks classifier for detection network intrusion. The proposed approach was tested on the KDD'99 Cup dataset, and obtained a promising result and outperformed the compared method for feature selection.

Recently, research conducted by Gu et al. [38] introduced an IDS based on SVM with the tabu-artificial bee colony for feature selection and parameter optimization simultaneously. The main contributions of their work included the adopting of the tabu search algorithm to improve the neighborhood search of ABC, so that it could speed up the convergence and prevent getting stuck in the local optimum. According to their experiments, although the accuracy rate was high, 94.53%, the false alarm rate was 7.028%.

Table 3 abstracts some characteristics of the discussed network intrusion detection approaches, concerning the focus areas, techniques, features, the datasets, and validation metrics used to evaluate

the models' performances. For the validation metrics, we present the best performing method in the paper. For the multi-task model, we present the evaluation measures of all tasks, if they exist, given by the "/" symbol. The acronyms for this table are described in Table 4.

Table 3. Selected literature focusing on network intrusion detection.

References	Year	Focus	Tech.	Anomaly Types	Dataset	Validation Metrics
[23]	2017	intrusion detection	SVM, K-means	DoS, Probe, U2R, R2L	KDD'99	ACC: 95.75%, FPR: 1.87%
[25]	2017	intrusion detection	NN with random weights	DoS, Probe, R2L, U2R	NSL-KDD	ACC: 84.12%
[28]	2017	intrusion detection	ACO, DT	DoS, Probe, R2L, U2R	NSL-KDD	ACC: 65%, FPR: 0%
[30]	2017	intrusion detection	PSO, KNN	DoS, Probe, R2L, U2R	KDD'99	ACC: - Dos: 99.91% - Probe: 94.41% - U2L: 99.77% - R2L: 99.73%
[24]	2018	intrusion detection	LS-SVM	DoS, Probe, U2R, and R2L	KDD'99	ACC: Over 99.6%
[26]	2018	intrusion detection	DAE, RF	DoS, Probe, R2L, U2R	KDD'99, NSL-KDD	Average ACC: 85.42% - 97.85%
[27]	2018	Anomaly detection	Fuzzy logic, GA	DoS, DDoS, Flash crowd	University dataset	ACC: 96.53%, FPR: 0.56%
[31]	2018	Intrusion detection	PSO, FLN	DoS, Probe, R2L, U2R	KDD'99	ACC: - Dos: 98.37% - Probe: 90.77% - U2L: 93.63% - R2L: 63.64%
[33]	2018	Anomaly detection	CSO, K-means	DoS, Probe, R2L, U2R	UCI-ML, NSL-KDD	ACC: 97.77%, FPR: 1.297%
[34]	2018	Intrusion detection & classification	ABC, AFS	DoS, Probe, R2L, U2R, Fuzzers, Analysis, Exploits, Generic, Worms, RA, Shellcode, Backdoors	NSL-KDD, UNSW-NB15	ACC: 97.5%, FPR: 0.01%
[32]	2019	anomaly detection	PSO, SVM, K-means, AFS	DoS,Probe, R2L, U2R, RA, RI, CI	KDD'99, Gas Pipeline	ACC: 95%
[35]	2019	Anomaly detection	GWO, CNN	DoS, Probe, U2R, R2L	DARPA'98, KDD'99	ACC: 97.92%/98.42% FPR: 3.6%/2.22%
[36]	2019	anomaly &misuse detection	Spark ML, LSTM	DSoS, DoS, Botnet, Brute Force SSH	ISCX-UNB	ACC: 97.29% FPR: 0.71%
[37]	2019	Anomaly detection	FA, C4.5, Bayesian Networks	DoS, Probe, U2R, R2L	KDD'99	DoS(ACC: 99.98%, FPR: 0.01%) Probe(ACC: 93.92%, FPR: 0.01%), R2L(ACC: 98.73%, FPR: 0%), U2R(ACC: 68.97%, FPR: 0%)
[38]	2019	Intrusion dectection	Tabu search, ABC, SVM	DoS, Probe, U2R, R2L	KDD'99	ACC: 94.53%, FPR: 7.028%

Table 4. The acronyms used in Table 3.

ACC: Accuracy	GA: Genetic Algorithms
FPR: False positive rate	CSO: Cuckoo Search Optimization
SVM: Support vector machine	ABC: Artificial bee colony
DT: Decision tree	AFS: Artificial fish swarm
NN: Neural Network	FA: Firefly algorithm
CNN: Convolutional neural network	GWO: Grey wolf optimization
KNN: K-nearest neighbors	Dos: Denial of Service
LS-SVM: Least squares support vector machines	R2L: Remote to local
DAE: Deep Auto-Encoder	U2R: User to Root
FLN: Fast learning network	RI:Response Injection
RF: Random forest	RA: Reconnaissance Attacks
ACO: Ant colony optimization	CI: Command Injection
PSO: Particle swarm optimization	

5.3. Phishing and SPAM Detection

A phishing attack is a cyber-attack that attempts to a steal user's identity or financial credentials. Today, phishing attacks are one of the most menacing threats on the Internet. Various novel intelligent approaches were used to cope with these problems.

The authors in [39] presented a phishing detection scheme called phishing email detection system (PEDS), which joined the evolving neural network and reinforcement learning. Their model obtained a 98.6% accuracy rate and a 1.8% false positive rate.

The authors in [40] introduced an anti-phishing method, which utilized several different ML algorithms and nineteen features to distinguish phishing websites from legitimate ones. The authors claimed that their model achieved a 99.39% true positive rate.

Another approach by Feng et al. [41], applied a neural network for identification the phishing web sites by adopting the Monte Carlo algorithm and risk minimization principle. Empirical results showed that their model reached a 97.71% precise detection rate and a 1.7% false alarm rate.

A recent study conducted by [42] introduced a real-time anti-phishing system, which utilized seven different classification algorithms and natural language processing (NLP) based features. According to the authors, their approach obtained a promising result with a 97.98% accuracy rate.

Another study [43] built a stacking model by combining GBDT, XGBoost, and LightGBM using URL and HTML features for classifying the phishing web pages. The authors reported that their approach reached a 98.60% accuracy rate.

The terminology "SPAM" refers to unsolicited bulk email (junk email). Spam email may lead to security issues and inappropriate contents. To overcome the drawbacks of this cyber-threats, recently scientists applied various novel, intelligent techniques to build spam filter systems.

Feng et al. [44] combined support machine vector and Naive Bayes to develop a spam filtering system. The proposed system was evaluated by the DATAMALL dataset and obtained a great spam-detection accuracy.

The authors in [45], designed a spam categorization technique using a modified cuckoo search to enhance the spam classification. In their work, the step size-cuckoo search was utilized for feature extraction, and the SVM was used for classification. The proposed approach was tested on two spam datasets—Bare-ling and Lemm-ling—and obtained a competitive result.

Later, research conducted by [46] proposed a system to filter the spam messages of Facebook using an SI-based and machine learning technique. The PSO algorithm was adopted for feature selection, and the SVM and decision tree for classification. The authors claimed that the proposed system was efficient. Unfortunately, the details of the results were not provided.

Recently, Aswani et al. [47] provided a hybrid approach for detecting the spam profiles on Twitter using social media analytics and bio-inspired computing. Specifically, they utilized a modified k-means-integrated levy flight firefly algorithm (LFA) with chaotic maps to identify spammers. A total of 14,235 profiles was used to evaluate the performance of the method. The empirical result showed that the proposed model was efficient with an accuracy of 97.98%.

A recent study conducted by Faris et al. [48] presented an email spam detection and identification system based on a genetic algorithm (GA) and a random weight network (RWN). According to the experiments, the proposed system obtained remarkable results in terms of accuracy, precision, and recall.

Table 5 exhibits some characteristics of the discussed phishing and spam detection approaches, concerning the focus areas, techniques, features, datasets, and validation metrics used to evaluate the models' performances. For the validation metrics, we present the best performing method in the paper. For the multi-task model, we present the evaluation measures of all tasks, if they exist, given by the "/" symbol. The acronyms for this table are described in Table 6.

Table 5. Selected literature focusing on phishing and spam identification.

Reference	Year	Focus	Tech.	Features	Dataset	Validation Metrics
[44]	2016	Spam detection	Naive Bayes, SVM	99 features	DATAMALL	Not provide
[45]	2017	Spam classification	CSO, SVM	101 features	Ling-spam corpus	ACC: 87%/88%
[39]	2018	Mail phishing detection	NN, RL	50 features	Self collection: 9900 samples	ACC: 98.6%, FPR: 1.8%
[40]	2018	Website phishing detection	RF, SVM, NN, logistic regression, naïve Bayes	19 features	Phishtank, Openphish, Alexa, Payment gateway, Top banking website	ACC: 99.09%
[41]	2018	Website phishing detection	NN	30 features	UCI repository phishing dataset	ACC: 97.71%, FPR: 1.7%.
[46]	2018	Spam message detection	PSO, DE, DT DB index, SVM,	13 features	Self collection: 200,000 samples	Not provide
[47]	2018	Spammer detection	LFA, FCM	21 features	Self collections: 14,235 samples	ACC: 97.98%
[42]	2019	Website phishing detection	Naive Bayes, KNN, Adaboost, K-star, SMO, RF, DT	104 features	Self collection: 73,575 samples	ACC: 97.98%
[43]	2019	Website phishing detection	GBDT, XGBoost, LightGBM	20 features	Self collection: - 1st: 49,947 samples - 2nd: 53,103 samples	ACC: 97.30%/98.60% FPR: 1.61%/1.24%
[48]	2019	spam detection	GA, RWN	140 features	Spam Assassin, LingSpam, CSDMC2010	ACC: 96.7%/93%/90.8%

Table 6. The acronyms used in Table 5.

ACC: Accuracy	GDBT: Gradient Boosting Decision Tree
FPR: False positive rate	RWN: Random Weight Network
SVM: Support vector machine	FCM: Fuzzy C-Means
DT: Decision tree	PSO: Particle swarm op timization
NN: Neural Network	GA: Genetic Algorithms
KNN: K-nearest neighbors	CSO: Cuckoo Search Optimization
RF: Random forest	LFA: Levy Flight Firefly Algorithm

5.4. Other: Counter APTs and Identify DGAs

In this part, the authors presents some existing works that leverage AI approaches to mitigate other types of cyber-threats. More precisely, the counter methods against APT attack and DGAs are described as follows.

5.4.1. Countering an Advanced, Persistent Threat

An advanced persistent threat (APT) is a sophisticated cyber-attack that uses advanced techniques to exploit sensitive data and remains undetected. The attackers often focus on valuable targets, such as large corporation's security agencies and government organizations, with the ultimate goal of long—term information stealing. To defend against APT attacks, scholars proposed a variety of AI techniques to deal with these cyber-threats.

In [49], the authors applied a decision tree to build IDS to detect APT attacks. It can detect intrusion from the beginning and quickly react to APT to minimize damage. Empirical results showed that the proposed system achieved a high rate of APT detection. Meanwhile, Sharma et al. [50] presented a framework architecture for the detection of APTs, which was based on multiple parallel classifiers. According to the authors, the proposed framework achieved great effectiveness and accuracy.

The authors in [51] investigated how deep neural networks (DNN), which used raw features of dynamic analysis could be employed for nation-state APT attribution. During evaluation with the training set containing 3200 samples, the proposed approach reached an accuracy of 94.6%.

Burnap et al. [52] used machine activity metrics and a self-organizing feature map approach to distinguish legitimate and malicious software. The authors reported that their method showed promise for APT detection.

Another approach [53] introduced a ML-based approach named MLAPT to identify and predict APTs. According to the authors, their system had the ability of early prediction of APT attacks. The experiments showed that MLAPT had a true positive rate and a false positive rate of 81.8% and 4.5% respectively.

5.4.2. Identifying Domain Names Generated by DGAs

Domain generation algorithms (DGAs) are algorithms that are used to create an immense number of pseudo-random domain names to hide the operator's command and control (C&C) server and evade detection. Lison et al. [54] adopted recurrent neural networks (RNN) to identify domain names generated by DGAs with high precision. According to the authors, the model could detect 97.3% of malware-generated domain names with a low false positive rate. Curtin et al. [55] also took a similar approach using the generalized likelihood ratio test (GLRT) and achieved promising results.

Yu et al. [56] performed a comparative analysis on convolutional neural network (CNN) and recurrent neural network (RNN) based architectures, tested using a dataset with one million domain names. The authors reported that all comparative models performed well with high accuracy rates and low false positive rates.

The authors in [57] introduced a novel long short-term memory network (LSTM) based algorithm to handle the multiclass imbalance problem in DGA malware detection. Based on the results obtained, the proposed algorithm provided an improvement as compared to the original LSTM.

In a recent study [58], the authors utilized IF-TF for a DGA and DNS covert channel detection system based machine learning. According to the authors, the proposed approach achieved outstanding accuracy at 99.92%.

Another approach in [59], proposed a framework for identification word-based DGAs by utilizing the frequency distribution of the words and an ensemble classifier constructed from naive Bayes, extra-trees, and logistic regression. The authors reported that their method outperformed the comparable ones.

Table 7 describes the main details of the selected studies focusing on APTs detection and identifying domains generated by DFGAs, concerning the focus areas, algorithms, datasets, and evaluation measures. For the validation metrics, we have presented the best model in the paper. List of acronyms used in this table is given in Table 8.

Table 7. Selected studies focusing on APT and DGA domains detection.

References	Year	Focus Area	Tech.	Features	Dataset	Validation Metrics
[49]	2017	APTs detection	DT	API calls	Self collection: 130 samples	ACC: 84.7%
[50]	2017	APTs detection	GT, DP, CART, SVM	Log events	Self collection	ACC: 98.5%, FPR: 2.4%
[51]	2017	nation-states APTs detection	DNN	Raw text	Self collection: 3200 samples	ACC: 94.6%
[54]	2017	DGA domains detection	RNN	Letter combinations	Self collection: over 2.9 million samples	ACC: 97.3%
[52]	2018	APTs detection	SOFM, DT, Bayesian network, SVM, NN	Machine activity metrics	Self collection: 1188 samples	ACC: 93.76%
[53]	2018	APTs detection and prediction	DT, KNN, SVM, EL	Network traffic	Self collection, university live traffic	ACC: 84.8%, FPR: 4.5%
[55]	2018	DGA domains detection	RNN	Characters	Self collection: 2.3 million samples	FPR: <=1%
[56]	2018	DGA domains detection	RNN, CNN	Strings	Self collection: 2 million samples	ACC: 97–98%
[57]	2018	DGA botnet detection	LSTM	Characters	Alexa, OSINT	F1:98.45%
[59]	2019	DGA detection	Ensemble classifier	words	Self collection: 1 million samples	ACC: 67.98%/89.91%/91.48%
[58]	2019	DGA, DNS covert chanel detection	TF-IDF	Strings	Self collection: 1 million samples	ACC: 99.92%

Table 8. Glossary of acronyms used for Table 7.

ACC: Accuracy	NN: Neural Networks
FPR: False positive rate	KNN: k-nearest neighbors
SVM: Support vector machine	EL: Ensemble learning
DT:Decision tree	RNN: Recurrent neural network
GP: Genetic programming	CNN: Convolutional neural network
DT: decision tree	TF - IDF: term frequency - inverse document frequency
CART: Classification and regression trees	LSTM: Long Short-Term Memory network
DBG-Model: Dynamic Bayesian game model	SOFM: Self Organising Feature Map
DNN: Deep neural network	SVM: Support Vector Machines

6. The Nefarious Use of AI

Regarding the fact that AI tools are already being developed open source, it is logical to expect that AI technologies may be leveraged for creating new types of advanced and sophisticated threats. In this section, we illustrate a range of feasible uses toward which AI could be put for nefarious ends. Some of them are already occurring in a limited form in practice but could be scaled up or strengthen with further technological advances in the future [60]. Figure 2 highlights some branches of leveraging of AI for malicious activities.

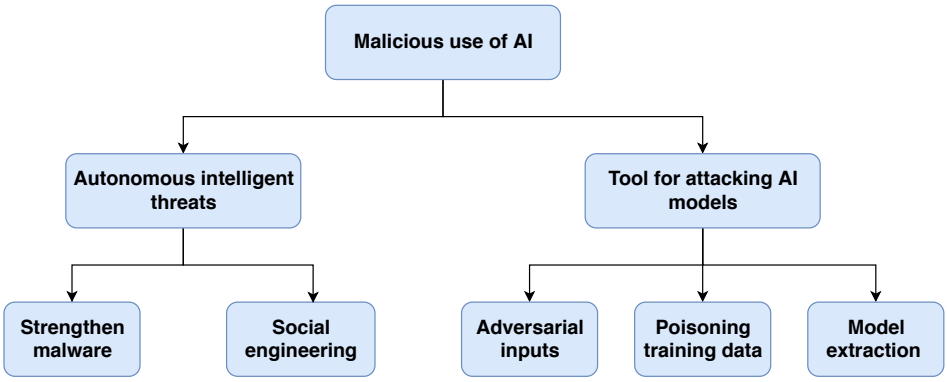

Figure 2. The use of AI for malicious activities in cybersecurity.

6.1. AI and Autonomy Intelligent Threats

The dangers of AI-enable threats appear with the ability to automate human abilities and processes and transcend current human capabilities. With the assistance of AI techniques, threat actors could enhance their weapons to make them more independent, complex, and hard to identify. In this section, the authors discuss how malicious actors can employ intelligent autonomous threats to endanger the defense system.

6.1.1. AI-Powered Malware

AI technologies can further be weaponized to increase the effectiveness of the malware, making it more autonomous, more sophisticated, faster, and harder to detect. With the support of AI, the new generation of malware becomes smarter and capable of operating autonomously. The intelligent malicious programs can self-propagate in a network or computer system based on a sequence of autonomous decisions, intelligently custom-made to the parameters of the host system, and autonomous malware capable of choosing the lateral movement techniques, thereby increasing the likelihood of fully compromising the targeted networks.

What is more, malware authors could adopt the ability to adapt to a new environment or to use the knowledge acquired from past occurrences of AI in creating intelligent viruses and malware or modeling adaptable attacks. Consequently, malware becomes independent, integrating into its environment, taking countermeasures against security tools and could leverage data acquired from the past to attack the system.

One of the ultimate goals of malware is to hide its presence and malicious intent to avoid being detected by anti-malware solutions. Cybercriminals will certainly discover ways to implement the most advanced technology into evasive techniques.

The researchers from IBM [61] presented malware enhanced by the DL technique that was capable of leveraging facial recognition, voice recognition, and geolocation to identify its target before for attacking.

In [62] Rigaki and Garcia adopted DL techniques to generate malicious malware samples that could avoid detection by simulating the behaviors of legitimate applications.

Concurrently to the development of malware, there are attempts being made to apply bio-inspired techniques into malware. For instance, Ney ea al. [63] presented how to compromise a computer by encoding malware into a DNA sequence. Later, the authors in [64] outlined a hypothetical swarm malware as a background for a future anti-malware system. More precisely, the swarm virus prototype simulated a swarm system behavior, and its information was stored and visualized in the form of a complex network. As a further improvement, the authors in [65] fused swarm based intelligence, a neural network, and a classical computer virus to form a neural swarm virus.

6.1.2. AI Used in Social Engineering Attacks

AI can be leveraged to mine large amounts of big datasets containing social network data to extract personally identifiable information, which can be used for compromising user accounts. What is more, based on user information, malicious actors could adopt AI to generate custom malicious links or create personalized phishing emails automatically.

There have been studies on adopting AI to carry out complex social engineering attacks. In [66,67], the authors introduced a long short-term memory (LSTM) neural network that was trained on social media posts to manipulate users into clicking on deceptive URLs.

6.2. AI as a Tool for Attacking AI Models

As AI is being integrated into security solutions, cybercriminals attempt to exploit vulnerabilities in this domain. Attacks on AI systems are typically discussed in the context of adversarial machine learning. The offenses on AI systems often appeared in three areas:

- Adversarial inputs: This is a technique where malicious actors design the inputs to make models predict erroneously in order to evade detection. Recent studies demonstrated how to generate adversarial malware samples to avoid detection. The authors of [68,69] crafted adversarial examples to attack the Android malware detection model. Meanwhile, scholars in [70] presented a generative adversarial network (GAN) based algorithm called MalGAN to craft adversarial samples, which was capable of bypassing black-box machine learning-based detection models. Another approach by Anderson et al. [71] adopted GAN to create adversarial domain names to avoid the detection of domain generation algorithms. The authors in [72] investigated adversarial generated methods to avoid detection by DL models. Meanwhile, in [73], the authors presented a framework based on reinforcement learning for attacking static portable executable (PE) anti-malware engines.
- Poisoning training data: In this kind of attack, the malicious actors could pollute the training data from which the algorithm was learning in such a way that reduced the detection capabilities of the system. Different domains are vulnerable to poisoning attacks; for example, network intrusion, spam filtering, or malware analysis [74,75].
- Model extraction attacks: These techniques are used to reconstruct the detection models or recover training data via black-box examination [76]. On this occasion, the attacker learns how ML algorithms work by reversing techniques. From this knowledge, the malicious actors know what the detector engines are looking for and how to avoid it.

Table 9 describes the main details of the selected studies focusing on malicious use of AI, with regard to the focus area, the techniques, the innovation point, and the main idea. The acronyms of this table are given in Table 10.

Table 9. Selected references in term of the malicious use of AI.

References	Year	Focus	Tech.	Innovation Point	Main Idea
[71]	2016	Adversarial attacks	GAN	New attack model	create adversarial domain names to avoid the detection of domain generation algorithms
[76]	2016	Stealing model	AE	model extraction attacks	extract target ML models by the machine learning prediction APIs
[66]	2016	Social engineering attacks	RNN	New attack model	Automated spear phishing campaign generator for social network
[63]	2017	Compromise computer	Encoding DNAs	Encoding malware to DNAs	compromise the computer by encoding malware in a DNA sequence
[68]	2017	Adversarial attacks	AE	New attack algorithm	adversarial attacks against deep learning based Android malware classification
[69]	2017	Adversarial attacks	AE	New attack algorithm	use the adversarial examples method to conduct new malware variants for malware detectors
[70]	2017	Adversarial attacks	GAN	New attack model	present a GAN based algorithm to craft malware that capable to bypass black-box machine learning-based detection models
[61]	2018	Malware creation	DNN	AI-powered malware	Leverage deep neural network enhance malware, make it more evasive and high targeting
[62]	2018	Malware creation	GAN	AI-powered malware	avoid detection by simulating the behaviors of legitimate applications
[64]	2018	Malware creation	ACO	SI-based malware	use ACO algorithms to create a prototype malware that have a decentralize behavior
[73]	2018	Adversarial attacks	AL	New attack method	a generic black-box for attacking static portable executable machine learning malware models
[72]	2018	Adversarial attacks	AM	New attack algorithm	adversarial generated methods to attack neural network-based malware detection
[74]	2018	Poisoning attack	EPD	New poisoning data method	present a novel poisoning approach that attack against machine learning algorithms used in IDSs
[75]	2018	Poisoning attack	AM	Analysis poisoning data method	present three kind of poisoning attacks on machine learning-based mobile malware detection
[67]	2018	Social engineering attacks	LSTM	New attack model	introduced a machine learning method to manipulate users into clicking on deceptive URLs
[65]	2019	Malware creation	ANN	next generation malware	fuse swarm base intelligence, neural network to form a new kind of malware

Table 10. The acronyms used in Table 9.

GAN: Generative adversarial network	EPD: Edge pattern detection
AE: Adversarial Examples	RNN: Recurrent neural network
SI: Swarm Intelligent	DNN: Deep neural network
ANN: Artificial Neural Network	ACO: Ant colony optimization
RL: Reinforcement learning	AM: Adversarial machine learning
AMB: Adversarial Malware Binaries	LSTM: Long short term memory

7. Challenges and Open Research Directions

In this section, we discuss the challenges when adopting AI-based approaches in practice. Additionally, we also offer a vision about some areas that need to be further research.

7.1. Challenges

AI methods have played a crucial role in cybersecurity applications and will continue in a promising direction that attracts investigations. However, some issues must be considered when applying AI-based techniques in cybersecurity. First, the accuracy of AI models is a significant barrier.

Specifically, false alarms can waste processing time, or an AI system might miss a cyberattack entirely. Another barrier to adoption is that many of the approaches proposed today are model-free methods. These models require a large quantity of training data, which are hard to obtain in real cybersecurity practice. Next, in designing AI-based solutions for cybersecurity, approaches need to consider the adversary. Adversarial attacks are hard to detect, prevent, and counter against as they are part of a battle between AI systems.

AI can help protect the system against cyber-threats but can also facilitate dangerous attacks; i.e., AI-based attacks. Malicious actors can leverage AI to make attacks flexible and more sophisticated to bypass detection methods to penetrate computer systems or networks.

7.2. Open Research Directions

There are diverse promising and open topics for incorporating AI techniques and cybersecurity. Some research areas are as follows.

First, the combination of several AI-based techniques in a defense solution may still an interesting research direction. For example, the incorporation of bio-inspired computation and ML/DL approaches shows promising results in malware detection [18–22] or [36–38] for detecting the network intrusion. Hence, the combination of these two techniques is a very potential research direction due to the number of bio-inspired algorithms exploited in cybersecurity still being limited.

Second, the corporation between a human intellect and machines for cyber defense also needs study. In this human–machine model, the agents will autonomously execute the task whilst humans can supervise and intervene only when necessary.

Third, there is literature proving that the threat actors could utilize the AI-based method to bypass or attack the AI models, such as in [68–72,75–77]. Hence, the defense strategy against these types of attacks would be an inevitable trend in the future.

Another aspect that necessitates being studied is the use of AI in malware, such as in [61,64,65]. Specifically, the combination of swarm communication and other AI-based techniques. Such malware will exhibit extremely high robustness of information preservation against swarm network damage. Swarm communication also exposes the research direction to apply this idea to other malware, such as worms, trojans, or ransomware so that their activities can be more distributed and stealth.

8. Discussion

The utilization of AI in cybersecurity creates new frontiers for security investigations. Scientists view AI as an essential response to the continuous growth in the number of and the increase in the complexity of cyber-threats, and the need for a quick reaction and substantially automatic responses to security attacks. On the other hand, AI technology also leads to some security issues that need to be resolved. In this section, we summarize the essential points in this study. Other methods to enhance cybersecurity are also mentioned. To conclude, the authors compared this study with several existing surveys.

It is clear from the literature that AI-based approaches could be adopted in the cyber domain, encompassing a variety of methods that have developed over many decades, have demonstrated effectiveness, and are currently in use.

At present, the prime targets for AI applications are malware classification and analysis, intrusion detection (focusing on anomaly network-based attacks), phishing and spam, and advanced persistent threat detection and characterization. Furthermore, a rapidly emerging topic for application is automated vulnerability testing and intrusion resistance.

Intrusion detection systems typically rely on hybridization techniques that combine several methods: signature-based methods for rapid detection of known threats with low false alarm rates and anomaly-based methods to flag deviations. What is more, another trend is combining with other computational intelligent models, such as ACO and PSO.

The absence of datasets for research and development in network intrusion is a problem. Precisely, publicly available datasets are extremely dated, such as DARPA (1998), KDD (1999), and NSL-KDD (2009), and the characteristics and volume of attacks have significantly changed since that time. What is more, the majority use of these datasets may offer a one-sided vision about collected data and not reflect real-world situations.

There are indications that AI-based models can be bypassed. Several published examples in the cybersecurity field indicate that the AI system can be challenged with the adversarial inputs or poisoning the training data. Furthermore, the potential threats of malicious use of AI need to be taken into account. For example, AI technology can be utilized to power malware, establish a spear-phishing campaign, or perform a social engineering attack.

Besides those previously mentioned topics, other research directions to enhance cybersecurity were also paid attention. For instance, in [78] the authors conducted a survey about the use of Kolmogorov complexity in the security and privacy domains. The adoption of these technologies in the cybersecurity realm was inspired by the feature free nature, and the absence of a need to tune the parameters.

In this work, we reviewed different AI techniques and methods used in defending against cyber-threats attacks and offered a vision of malicious use of AI technology as potential threats. In order to ensure the novelty and new contribution of our survey, we thoroughly compared our work with existing surveys, as shown in Table 11.

Table 11. A comparison between our surveys and existing surveys in the literature.

Content		**References**							
Year		[7]	[8]	[6]	[3]	[4]	[5]	[2]	**This Survey**
		2018	2018	2018	2018	2018	2019	2019	2019
AI methods	Machine learning	x	x		x	x		x	x
	Deep learning	x	x	x		x	x		x
	Bio-inspire computing							x	x
Defense applications	Malware detection	x	x	x	x	x	x	x	x
	Intrusion detection	x	x	x	x	x	x		x
	Phishing detection	x	x		x		x	x	x
	Spam identification	x			x		x	x	x
	APTs detection			x	x			x	x
	DGAs detection	x			x		x		x
Malicous use of AI	Ai-powerd malware							x	x
	Attack against AI	x		x				x	x
	Social engineering attacks							x	x

9. Conclusions

Dramatic advances in information technology have led to the emergence of new challenges for cybersecurity. The computational complexity of cyber-attacks requires new approaches which are more robust, scalable, and flexible. This article focuses on the application of the AI-based technique in cybersecurity issues. Specifically, we present the application of AI in malware detection, intrusion detection, APT, and other domains, such as spam detection and phishing detection. Furthermore, our manuscript offers a vision of how AI could be adopted for malicious use.

In contemporary research, the primary targets for AI application in cybersecurity are network intrusion detection, malware analysis and classification, phishing, and spam emails. In those areas, the adoption of DL gradually became the primary trend. Furthermore, the combination of other intelligent techniques, such as bio-inspired methods, together with ML/DL, also attracted the attention of researchers. Such combinations yield very promising results and continue a trend for further research.

Although the role of AI in resolving cybersecurity matters continues to be researched, some of the problems that exist around the deployment of AI-based defenses are also striking. For instance,

the adversarial attack against the AI models or the emergence of autonomous intelligent malware. Hence, research on discovering solutions to these threats should be further explored.

Author Contributions: All the authors are responsible for the concept of the paper, the results presented, and the writing, and contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The following grants are acknowledged for the financial support provided for this research: grant of SGS, number SP2020/78, VSB Technical University of Ostrava.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial intelligence
ML	Machine learning
DL	Deep learning
DT	Decision trees
SVM	Support vector machines
KNN	K-nearest neighbor
RF	Random forest
AR	Association rule algorithms
EL	Ensemble learning
PCA	Principal component analysis
FNN	Feedforward neural networks
CNNs	Convolutional neural networks
RNN	Recurrent neural networks
DBNs	Deep belief networks
SAE	Stacked autoencoders
GANs	Generative adversarial networks
RBMs	Restricted Boltzmann machines
EDLNs	Ensemble of deep learning networks
GA	Genetic algorithms
ES	Evolution strategies
ACO	Ant colony optimization
PSO	Particle swarm optimization
AIS	Artificial immune systems (AIS)

References

1. Buczak, A.L.; Guven, E. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Commun. Surv. Tutor.* **2015**, *18*, 1153–1176. [CrossRef]
2. Torres, J.M.; Comesaña, C.I.; García-Nieto, P.J. Machine learning techniques applied to cybersecurity. *Int. J. Mach. Learn. Cybern.* **2019**, 1–14. [CrossRef]
3. Guan, Z.; Bian, L.; Shang, T.; Liu, J. When machine learning meets security issues: A survey. In Proceedings of the 2018 IEEE International Conference on Intelligence and Safety for Robotics (ISR), Shenyang, China, 24–27 August 2018; pp. 158–165.
4. Xin, Y.; Kong, L.; Liu, Z.; Chen, Y.; Li, Y.; Zhu, H.; Gao, M.; Hou, H.; Wang, C. Machine learning and deep learning methods for cybersecurity. *IEEE Access* **2018**, *6*, 35365–35381. [CrossRef]
5. Berman, D.S.; Buczak, A.L.; Chavis, J.S.; Corbett, C.L. A survey of deep learning methods for cyber security. *Information* **2019**, *10*, 122. [CrossRef]
6. Wickramasinghe, C.S.; Marino, D.L.; Amarasinghe, K.; Manic, M. Generalization of Deep Learning for Cyber-Physical System Security: A Survey. In Proceedings of the IECON 2018—44th Annual Conference of the IEEE Industrial Electronics Society, Washington, DC, USA, 21–23 October 2018; pp. 745–751.
7. Apruzzese, G.; Colajanni, M.; Ferretti, L.; Guido, A.; Marchetti, M. On the effectiveness of machine and deep learning for cyber security. In Proceedings of the 2018 10th International Conference on Cyber Conflict (CyCon), Tallinn, Estonia, 30 May–1 June 2018; pp. 371–390.

8. Li, J.H. Cyber security meets artificial intelligence: A survey. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, 1462–1474. [CrossRef]
9. Xu, Z.; Ray, S.; Subramanyan, P.; Malik, S. Malware detection using machine learning based analysis of virtual memory access patterns. In Proceedings of the Conference on Design, Automation & Test in Europe, Lausanne, Switzerland, 27–31 March 2017; pp. 169–174.
10. Chowdhury, M.; Rahman, A.; Islam, R. Malware analysis and detection using data mining and machine learning classification. In Proceedings of the International Conference on Applications and Techniques in Cyber Security and Intelligence, Ningbo, China, 16–18 June 2017; pp. 266–274.
11. Hashemi, H.; Azmoodeh, A.; Hamzeh, A.; Hashemi, S. Graph embedding as a new approach for unknown malware detection. *J. Comput. Virol. Hacking Tech.* **2017**, *13*, 153–166. [CrossRef]
12. Ye, Y.; Chen, L.; Hou, S.; Hardy, W.; Li, X. DeepAM: A heterogeneous deep learning framework for intelligent malware detection. *Knowl. Inf. Syst.* **2018**, *54*, 265–285. [CrossRef]
13. McLaughlin, N.; Martinez del Rincon, J.; Kang, B.; Yerima, S.; Miller, P.; Sezer, S.; Safaei, Y.; Trickel, E.; Zhao, Z.; Doupé, A.; et al. Deep android malware detection. In Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy, Scottsdale, AZ, USA, 22–24 March 2017; pp. 301–308.
14. Li, J.; Sun, L.; Yan, Q.; Li, Z.; Srisa-an, W.; Ye, H. Significant permission identification for machine-learning-based android malware detection. *IEEE Trans. Ind. Inform.* **2018**, *14*, 3216–3225. [CrossRef]
15. Zhu, H.J.; You, Z.H.; Zhu, Z.X.; Shi, W.L.; Chen, X.; Cheng, L. DroidDet: Effective and robust detection of android malware using static analysis along with rotation forest model. *Neurocomputing* **2018**, *272*, 638–646. [CrossRef]
16. Karbab, E.B.; Debbabi, M.; Derhab, A.; Mouheb, D. MalDozer: Automatic framework for android malware detection using deep learning. *Digit. Investig.* **2018**, *24*, S48–S59. [CrossRef]
17. Wang, W.; Zhao, M.; Wang, J. Effective android malware detection with a hybrid model based on deep autoencoder and convolutional neural network. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 3035–3043. [CrossRef]
18. Ab Razak, M.F.; Anuar, N.B.; Othman, F.; Firdaus, A.; Afifi, F.; Salleh, R. Bio-inspired for features optimization and malware detection. *Arab. J. Sci. Eng.* **2018**, *43*, 6963–6979. [CrossRef]
19. Altaher, A.; Barukab, O.M. Intelligent Hybrid Approach for Android Malware Detection based on Permissions and API Calls. *Int. J. Adv. Comput. Sci. Appl.* **2017**, *8*, 60–67. [CrossRef]
20. Bhattacharya, A.; Goswami, R.T.; Mukherjee, K. A feature selection technique based on rough set and improvised PSO algorithm (PSORS-FS) for permission based detection of Android malwares. *Int. J. Mach. Learn. Cybern.* **2019**, *10*, 1893–1907. [CrossRef]
21. Alejandre, F.V.; Cortés, N.C.; Anaya, E.A. Feature selection to detect botnets using machine learning algorithms. In Proceedings of the 2017 International Conference on Electronics, Communications and Computers (CONIELECOMP), Cholula, Mexico, 22–24 February 2017; pp. 1–7.
22. Fatima, A.; Maurya, R.; Dutta, M.K.; Burget, R.; Masek, J. Android Malware Detection Using Genetic Algorithm based Optimized Feature Selection and Machine Learning. In Proceedings of the 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, 1–3 July 2019; pp. 220–223.
23. Al-Yaseen, W.L.; Othman, Z.A.; Nazri, M.Z.A. Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system. *Expert Syst. Appl.* **2017**, *67*, 296–303. [CrossRef]
24. Kabir, E.; Hu, J.; Wang, H.; Zhuo, G. A novel statistical technique for intrusion detection systems. *Future Gener. Comput. Syst.* **2018**, *79*, 303–318. [CrossRef]
25. Ashfaq, R.A.R.; Wang, X.Z.; Huang, J.Z.; Abbas, H.; He, Y.L. Fuzziness based semi-supervised learning approach for intrusion detection system. *Inf. Sci.* **2017**, *378*, 484–497. [CrossRef]
26. Shone, N.; Ngoc, T.N.; Phai, V.D.; Shi, Q. A deep learning approach to network intrusion detection. *IEEE Trans. Emerg. Top. Comput. Intell.* **2018**, *2*, 41–50. [CrossRef]
27. Hamamoto, A.H.; Carvalho, L.F.; Sampaio, L.D.H.; Abrão, T.; Proença, M.L., Jr. Network anomaly detection system using genetic algorithm and fuzzy logic. *Expert Syst. Appl.* **2018**, *92*, 390–402. [CrossRef]

28. Botes, F.H.; Leenen, L.; De La Harpe, R. Ant colony induced decision trees for intrusion detection. In Proceedings of the 16th European Conference on Cyber Warfare and Security, Dublin, Ireland, 29–30 June 2017; pp. 53–62.
29. Otero, F.E.; Freitas, A.A.; Johnson, C.G. Inducing decision trees with an ant colony optimization algorithm. *Appl. Soft Comput.* **2012**, *12*, 3615–3626. [CrossRef]
30. Syarif, A.R.; Gata, W. Intrusion detection system using hybrid binary PSO and K-nearest neighborhood algorithm. In Proceedings of the 2017 11th International Conference on Information & Communication Technology and System (ICTS), Surabaya, India, 31 October 2017; pp. 181–186.
31. Ali, M.H.; Al Mohammed, B.A.D.; Ismail, A.; Zolkipli, M.F. A new intrusion detection system based on fast learning network and particle swarm optimization. *IEEE Access* **2018**, *6*, 20255–20261. [CrossRef]
32. Chen, W.; Liu, T.; Tang, Y.; Xu, D. Multi-level adaptive coupled method for industrial control networks safety based on machine learning. *Saf. Sci.* **2019**, *120*, 268–275. [CrossRef]
33. Garg, S.; Batra, S. Fuzzified cuckoo based clustering technique for network anomaly detection. *Comput. Electr. Eng.* **2018**, *71*, 798–817. [CrossRef]
34. Hajisalem, V.; Babaie, S. A hybrid intrusion detection system based on ABC-AFS algorithm for misuse and anomaly detection. *Comput. Netw.* **2018**, *136*, 37–50. [CrossRef]
35. Garg, S.; Kaur, K.; Kumar, N.; Kaddoum, G.; Zomaya, A.Y.; Ranjan, R. A Hybrid Deep Learning based Model for Anomaly Detection in Cloud Datacentre Networks. *IEEE Trans. Netw. Serv. Manag.* **2019**. [CrossRef]
36. Khan, M.A.; Karim, M.; Kim, Y. A Scalable and Hybrid Intrusion Detection System Based on the Convolutional-LSTM Network. *Symmetry* **2019**, *11*, 583. [CrossRef]
37. Selvakumar, B.; Muneeswaran, K. Firefly algorithm based feature selection for network intrusion detection. *Comput. Secur.* **2019**, *81*, 148–155.
38. Gu, T.; Chen, H.; Chang, L.; Li, L. Intrusion detection system based on improved abc algorithm with tabu search. *IEEJ Trans. Electr. Electron. Eng.* **2019**, *14*. [CrossRef]
39. Smadi, S.; Aslam, N.; Zhang, L. Detection of online phishing email using dynamic evolving neural network based on reinforcement learning. *Decis. Support Syst.* **2018**, *107*, 88–102. [CrossRef]
40. Jain, A.K.; Gupta, B.B. Towards detection of phishing websites on client-side using machine learning based approach. *Telecommun. Syst.* **2018**, *68*, 687–700. [CrossRef]
41. Feng, F.; Zhou, Q.; Shen, Z.; Yang, X.; Han, L.; Wang, J. The application of a novel neural network in the detection of phishing websites. *J. Ambient. Intell. Humaniz. Comput.* **2018**, 1–15. [CrossRef]
42. Sahingoz, O.K.; Buber, E.; Demir, O.; Diri, B. Machine learning based phishing detection from URLs. *Expert Syst. Appl.* **2019**, *117*, 345–357. [CrossRef]
43. Li, Y.; Yang, Z.; Chen, X.; Yuan, H.; Liu, W. A stacking model using URL and HTML features for phishing webpage detection. *Future Gener. Comput. Syst.* **2019**, *94*, 27–39. [CrossRef]
44. Feng, W.; Sun, J.; Zhang, L.; Cao, C.; Yang, Q. A support vector machine based naive Bayes algorithm for spam filtering. In Proceedings of the 2016 IEEE 35th International Performance Computing and Communications Conference (IPCCC), Las Vegas, NV, USA, 9–11 December 2016; pp. 1–8.
45. Kumaresan, T.; Palanisamy, C. E-mail spam classification using S-cuckoo search and support vector machine. *Int. J. Bio-Inspired Comput.* **2017**, *9*, 142–156. [CrossRef]
46. Sohrabi, M.K.; Karimi, F. A feature selection approach to detect spam in the Facebook social network. *Arab. J. Sci. Eng.* **2018**, *43*, 949–958. [CrossRef]
47. Aswani, R.; Kar, A.K.; Ilavarasan, P.V. Detection of spammers in twitter marketing: A hybrid approach using social media analytics and bio inspired computing. *Inf. Syst. Front.* **2018**, *20*, 515–530. [CrossRef]
48. Faris, H.; Ala'M, A.Z.; Heidari, A.A.; Aljarah, I.; Mafarja, M.; Hassonah, M.A.; Fujita, H. An intelligent system for spam detection and identification of the most relevant features based on evolutionary random weight networks. *Inf. Fusion* **2019**, *48*, 67–83. [CrossRef]
49. Moon, D.; Im, H.; Kim, I.; Park, J.H. DTB-IDS: An intrusion detection system based on decision tree using behavior analysis for preventing APT attacks. *J. Supercomput.* **2017**, *73*, 2881–2895. [CrossRef]
50. Sharma, P.K.; Moon, S.Y.; Moon, D.; Park, J.H. DFA-AD: A distributed framework architecture for the detection of advanced persistent threats. *Clust. Comput.* **2017**, *20*, 597–609. [CrossRef]
51. Rosenberg, I.; Sicard, G.; David, E.O. DeepAPT: Nation-state APT attribution using end-to-end deep neural networks. In Proceedings of the International Conference on Artificial Neural Networks, Alghero, Sardinia, Italy, 11–14 September 2017; pp. 91–99.

52. Burnap, P.; French, R.; Turner, F.; Jones, K. Malware classification using self organising feature maps and machine activity data. *Comput. Secur.* **2018**, *73*, 399–410. [CrossRef]
53. Ghafir, I.; Hammoudeh, M.; Prenosil, V.; Han, L.; Hegarty, R.; Rabie, K.; Aparicio-Navarro, F.J. Detection of advanced persistent threat using machine-learning correlation analysis. *Future Gener. Comput. Syst.* **2018**, *89*, 349–359. [CrossRef]
54. Lison, P.; Mavroeidis, V. Automatic detection of malware-generated domains with recurrent neural models. *arXiv* **2017**, arXiv:1709.07102.
55. Curtin, R.R.; Gardner, A.B.; Grzonkowski, S.; Kleymenov, A.; Mosquera, A. Detecting DGA domains with recurrent neural networks and side information. *arXiv* **2018**, arXiv:1810.02023.
56. Yu, B.; Pan, J.; Hu, J.; Nascimento, A.; De Cock, M. Character level based detection of DGA domain names. In Proceedings of the IEEE 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–8.
57. Tran, D.; Mac, H.; Tong, V.; Tran, H.A.; Nguyen, L.G. A LSTM based framework for handling multiclass imbalance in DGA botnet detection. *Neurocomputing* **2018**, *275*, 2401–2413. [CrossRef]
58. Wang, Z.; Dong, H.; Chi, Y.; Zhang, J.; Yang, T.; Liu, Q. DGA and DNS Covert Channel Detection System based on Machine Learning. In Proceedings of the 3rd International Conference on Computer Science and Application Engineering, Sanya, China, 22–24 October 2019; p. 156.
59. Yang, L.; Zhai, J.; Liu, W.; Ji, X.; Bai, H.; Liu, G.; Dai, Y. Detecting Word-Based Algorithmically Generated Domains Using Semantic Analysis. *Symmetry* **2019**, *11*, 176. [CrossRef]
60. Thanh, C.T.; Zelinka, I. A Survey on Artificial Intelligence in Malware as Next-Generation Threats. *Mendel* **2019**, *25*, 27–34. [CrossRef]
61. Stoecklin, M.P. DeepLocker: How AI Can Power a Stealthy New Breed of Malware. *Secur. Intell.* **2018**, *8*. Available online: https://securityintelligence.com/deeplocker-how-ai-can-power-a-stealthy-new-breed-of-malware/ (accessed on 8 February 2020).
62. Rigaki, M.; Garcia, S. Bringing a gan to a knife-fight: Adapting malware communication to avoid detection. In Proceedings of the 2018 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 24 May 2018; pp. 70–75.
63. Ney, P.; Koscher, K.; Organick, L.; Ceze, L.; Kohno, T. Computer Security, Privacy, and DNA Sequencing: Compromising Computers with Synthesized DNA, Privacy Leaks, and More. In Proceedings of the 26th USENIX Security Symposium (USENIX Security 17), Vancouver, BC, Canada, 16–18 August 2017; pp. 765–779.
64. Zelinka, I.; Das, S.; Sikora, L.; Šenkeřík, R. Swarm virus-Next-generation virus and antivirus paradigm? *Swarm Evol. Comput.* **2018**, *43*, 207–224. [CrossRef]
65. Truong, T.C.; Zelinka, I.; Senkerik, R. Neural Swarm Virus. In *Swarm, Evolutionary, and Memetic Computing and Fuzzy and Neural Computing*; Springer: Berlin, Germany, 2019; pp. 122–134.
66. Seymour, J.; Tully, P. Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter. *Black Hat USA* **2016**, *37*.#t=00:17,00:54. [CrossRef]
67. Seymour, J.; Tully, P. Generative Models for Spear Phishing Posts on Social Media. *arXiv* **2018**, arXiv:1802.05196.
68. Grosse, K.; Papernot, N.; Manoharan, P.; Backes, M.; McDaniel, P. Adversarial examples for malware detection. In Proceedings of the European Symposium on Research in Computer Security, Oslo, Norway, 11–15 September 2017; pp. 62–79.
69. Yang, W.; Kong, D.; Xie, T.; Gunter, C.A. Malware detection in adversarial settings: Exploiting feature evolutions and confusions in android apps. In Proceedings of the 33rd Annual Computer Security Applications Conference, Orlando, FL, USA, 4–8 December 2017; pp. 288–302.
70. Hu, W.; Tan, Y. Generating adversarial malware examples for black-box attacks based on GAN. *arXiv* **2017**, arXiv:1702.05983.
71. Anderson, H.S.; Woodbridge, J.; Filar, B. DeepDGA: Adversarially-tuned domain generation and detection. In Proceedings of the 2016 ACM Workshop on Artificial Intelligence and Security, Vienna, Austria, 28 October 2016; pp. 13–21.
72. Kolosnjaji, B.; Demontis, A.; Biggio, B.; Maiorca, D.; Giacinto, G.; Eckert, C.; Roli, F. Adversarial malware binaries: Evading deep learning for malware detection in executables. In Proceedings of the 2018 26th European Signal Processing Conference (EUSIPCO), Rome, Italy, 3–7 September 2018; pp. 533–537.

73. Anderson, H.S.; Kharkar, A.; Filar, B.; Evans, D.; Roth, P. Learning to evade static PE machine learning malware models via reinforcement learning. *arXiv* **2018**, arXiv:1801.08917.
74. Li, P.; Liu, Q.; Zhao, W.; Wang, D.; Wang, S. BEBP: An poisoning method against machine learning based idss. *arXiv* **2018**, arXiv:1803.03965.
75. Chen, S.; Xue, M.; Fan, L.; Hao, S.; Xu, L.; Zhu, H.; Li, B. Automated poisoning attacks and defenses in malware detection systems: An adversarial machine learning approach. *Comput. Secur.* **2018**, *73*, 326–344. [CrossRef]
76. Tramèr, F.; Zhang, F.; Juels, A.; Reiter, M.K.; Ristenpart, T. Stealing machine learning models via prediction apis. In Proceedings of the 25th USENIX Security Symposium (USENIX Security 16), Austin, TX, USA, 10–12 August 2016; pp. 601–618.
77. Carlini, N.; Liu, C.; Erlingsson, Ú.; Kos, J.; Song, D. The secret sharer: Evaluating and testing unintended memorization in neural networks. In Proceedings of the 28th USENIX Security Symposium (USENIX Security 19), San Diego, California, CA, USA, 14–16 August 2019; pp. 267–284.
78. Resende, J.S.; Martins, R.; Antunes, L. A Survey on Using Kolmogorov Complexity in Cybersecurity. *Entropy* **2019**, *21*, 1196. [CrossRef]