applied sciences

MDPI

Review

A Systematic Review of Defensive and Offensive Cybersecurity with Machine Learning

Imatitikua D. Aiyanyo [1], Hamman Samuel [2,*] and Heuiseok Lim [1]

[1] Department of Computer Science and Engineering, College of Informatics, Korea University, Seoul 02841, Korea; titi@korea.ac.kr (I.D.A.); limhseok@korea.ac.kr (H.L.)
[2] Department of Computing Science, University of Alberta, Edmonton, AB T6G 2E8, Canada
* Correspondence: hwsamuel@ualberta.ca

Received: 10 August 2020; Accepted: 20 August 2020; Published: 22 August 2020

Abstract: This is a systematic review of over one hundred research papers about machine learning methods applied to defensive and offensive cybersecurity. In contrast to previous reviews, which focused on several fragments of research topics in this area, this paper systematically and comprehensively combines domain knowledge into a single review. Ultimately, this paper seeks to provide a base for researchers that wish to delve into the field of machine learning for cybersecurity. Our findings identify the frequently used machine learning methods within supervised, unsupervised, and semi-supervised machine learning, the most useful data sets for evaluating intrusion detection methods within supervised learning, and methods from machine learning that have shown promise in tackling various threats in defensive and offensive cybersecurity.

Keywords: cybersecurity; machine learning; artificial intelligence; data mining; defensive security; offensive security; intrusion detection systems

1. Introduction

In the fight against malicious threats, there has been collaborative support from experts to design different cyber defense systems. Both researchers and designers respectively have the same goals: maintain the privacy, integrity, and accessibility of information through the cyber defense systems against both internal and external threats. The main goal of cybersecurity systems is to combat security threats originating from online sources, including viruses, Trojans, worms, spam, and botnets [1]. These systems defend against cybersecurity threats at both the network and host levels. Network-based defense systems make use of the network flow while host-based defense systems control workstation's upcoming data by mechanisms designed in firewalls, antiviruses, and Intrusion Detection Systems (IDS).

These mechanisms monitor, track, and block viruses and other malicious cyberattacks. However, these methods do not completely eliminate vulnerabilities, threats, and attacks because the design and implementation of software and network infrastructure is inherently imperfect. The old adage that "security chain is only as strong as the weakest link" sums this up aptly, because a single weak spot within modern software and network infrastructures can lead to cascading security compromises at multiple sub-levels [2]. This has led to the constant cycle of patches to protect cyberspace infrastructure, but this has not deterred attackers. Thus, building defense systems for known attacks is insufficient in protecting users. Effective cybersecurity is more critical than ever, as modern attacks are being initiated with the intent for cyberwarfare by well-trained and well-funded militaries and criminal organizations. Moreover, the intensity of attacks has increased and correspondingly the impact of intrusions, as people and organizations get more connected via the Internet of Things (IoT) [3].

Advanced methods are needed to discover previously unknown cyber intrusions and techniques towards a more dependable cybersecurity infrastructure, including both defensive and offensive approaches. Defensive approaches use reactive strategies that focus on prevention, detection, and responses. This is the more traditional method to keep networks safe from cyber criminals, and requires a thorough understanding of the system to be secured. Preventive measures are developed from understanding of the system and potential weak points [4]. On the other hand, offensive approaches are counterpoint to defensive methods, and proactively predict and remove threats in the system using ethical hacking techniques. Security experts mimic exploits and attacks as cyber attackers would. Ultimately, experts aim to eliminate vulnerabilities by identifying them ahead of time [4]. Due to the accessibility of vast volumes of data and cyber criminals trying to gain illegal access to cyberinfrastructures, various Artificial Intelligence (AI) and Machine Learning (ML) techniques have been explored. This is because ML-based cybersecurity solutions, both offensive and defensive, can handle and analyze large amounts of data and complex detection logic where traditional methods would struggle.

Previous reviews in this area have focused on several fragments of research topics. This paper systematically combines the knowledge base in cybersecurity with ML. The goal of our research is to provide a baseline for readers, covering ML techniques, objectives, and effectiveness in cybersecurity, as well as current challenges and future directions of ML techniques in cybersecurity. We focus on a clear depiction of various ML methods, including Data Mining (DM) and computational intelligence. We survey nearly two decades of research papers on application of ML techniques to cybersecurity. We also review and contextualize the literature through the Six Dimensions of Intersection of AI/ML and Cybersecurity (AI-ML-CS) framework [3]. Our systematic review focuses on two dimensions: firstly data and new information frontiers and, secondly, algorithms for AI/ML and cybersecurity.

The paper is organized as follows, in line with the IMRAD (Introduction, Methods, Results, and Discussion) organization structure typically used in scientific literature [5]. After the Introduction section, we provide an overview of our Methodology grounded in Systematic Literature Reviews (SLR). The outcomes from implementation of the SLR methods are contained in the Results section, while applications of the results are covered in the Discussion section.

2. Methodology

There are other complementary surveys on the topic of ML and cyberattacks [6] and cybersecurity [7]. In order to add to domain knowledge in this area, this review aims at producing an impartial and comprehensive search of the resources considered from the defensive and offensive cybersecurity perspectives. This involves utilizing systematic methods and secondary data, while critically appraising research studies in order to synthesize findings qualitatively and quantitatively. The recent interest in the field of cybersecurity approaches to ML has not yet resulted in an effort to survey the underlying concepts, methods, and problems systematically. Our research is partially using the Kitchenham and Charters methodology for SLRs [8]. This includes three critical steps that pre-define a review protocol to reduce potential researcher bias: outlining the research questions, generating a search strategy, and specifying a selection criteria. Overall, this systematic review adheres to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [9]. To fulfill the requirements of PRISMA, this paper has been structured in accordance with the required sections, and we have provided the PRISMA flow diagram in Figure 1, as well as the PRISMA checklist with cross-references in Appendix A.

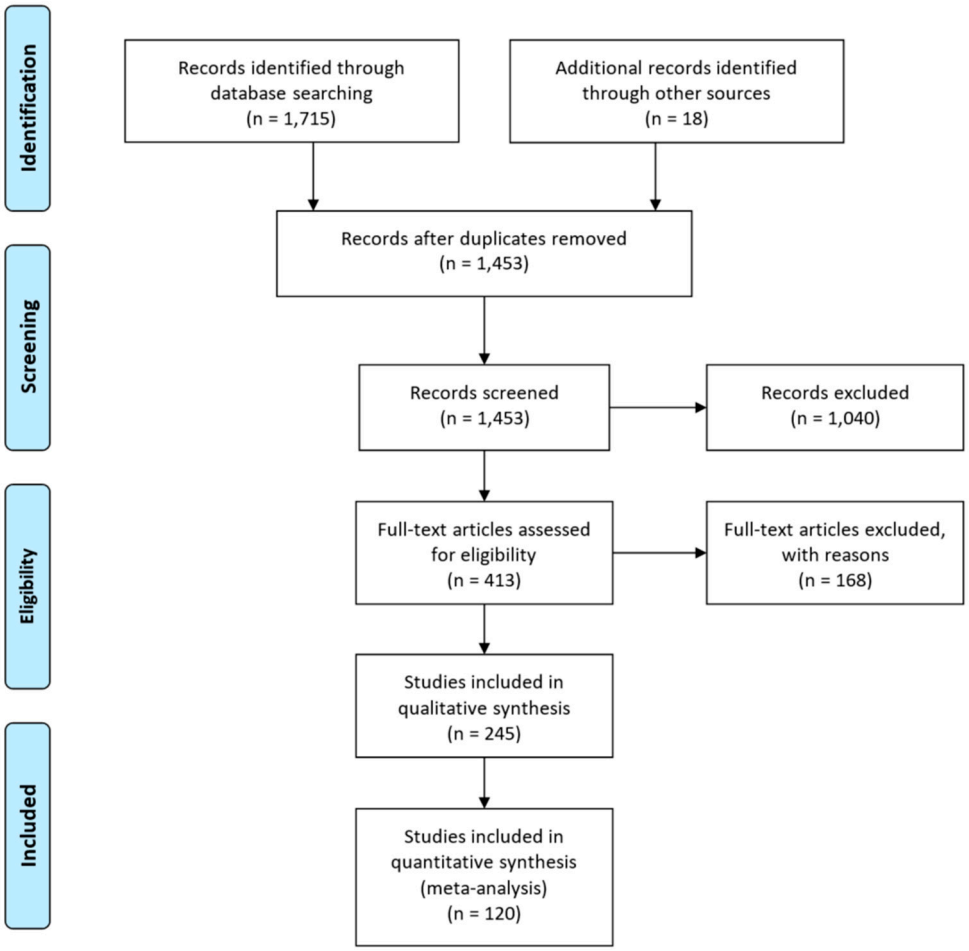

Figure 1. Search strategy steps as Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.

2.1. Research Questions

We cover the following three main research questions in this research that the systematic review is aiming to answer, with emphasis on current ML methods being used in cybersecurity. These questions are formulated to be relevant to both researchers and practitioners in cybersecurity. In line with the AI-ML-CS framework, our research questions explore the need and availability of training data sets, as well as the need for the two-pronged approach using passive (forensic or defensive) versus proactive (offensive) cybersecurity strategies in algorithms [3].

- What ML techniques have been used in offensive and defensive cybersecurity?
- What data sets are used in training supervised ML models by researchers?
- What cybersecurity threats can be better tackled by ML in cybersecurity?

2.2. Search Strategy

Figure 1 shows the study selection divided into three different stages: identification, screening, and eligibility. For the identification step, we shortlisted five major databases: ACM Digital Library, IEEE Xplore, Springer Link, Science Direct, and Scopus. Scanning bibliographies of relevant articles, we also searched for resources outside those databases. For example, Google Scholar was also investigated but exempted from our sources, due to overlapping citations with the other databases. Google Scholar is not a publishing entity, and mostly indexes citations and papers from other primary database sources. On the other hand, the selected sources are original publishers who index original

papers from self-managed conferences and journals by ACM, IEEE, Springer, and Elsevier. Hence, the overlap between publications retrieved from the selected databases is minimal. In order to collect all relevant existing research, we employed groups of several keywords. Despite major contributions to the cybersecurity industry starting from the 2010s, we chose our timeline based on the impact of ML in the early 2000s and included contributions made prior to 2010, as well.

In line with PRISMA, we also self-evaluate our sampling methods for any hidden bias, which could lead to diverging outcomes from systematic studies [10]. Our data collection strategy had no known hidden bias. The databases selected are well-known, and we were not limited by any issues regarding access to manuscripts. We have also taken into account citation counts and year of publication for the selected studies, as documented in Appendix B, and hence, it is expected that any potential errors from missing or incomplete data will be minimal.

After collection, we identified duplicates manually based on citation similarity. The following search terms were used (Box 1), with ML, artificial intelligence, and DM grouped inclusively to ensure the presence of either word would return the matched articles. To filter for only security-related articles, cybersecurity was used exclusively in the search query to return ML articles specific to cybersecurity.

Box 1. Search terms.

```
(machine learning OR artificial intelligence OR data mining) AND
cybersecurity
```

In the screening step, a team of researchers independently filtered the collection based on their titles and abstracts. Next, the researchers examined the full contents of the selected collection. The eligibility step reviewed articles using outlined search criteria to guarantee only relevant articles had been collected. In the event of differing opinions in relation to the credibility of any selected article, meetings were held to re-examine the full article until consensus was reached. Ultimately, 245 articles on cybersecurity approaches using ML techniques were collected. From these, further criteria were applied based on uniqueness of contributions to ML and cybersecurity, as well as removing duplicated contributions to get the final listing of 120 studies, listed with references in Appendix B. We also extracted metadata from the final selected set, including author's name, title, year of publication, publishing type, citation, data set, objectives, ML techniques, current and future challenges, offensive or defensive cybersecurity approach, and data sets used.

2.3. Search Criteria

A systematic search of the literature concerning offensive and defensive cybersecurity approaches using ML techniques was performed. Table 1 shows several criteria defined in order to find high-quality articles to answer our research questions. These criteria were applied in order to determine which articles would be included or excluded from all articles across each phase of the selection process.

Table 1. Search inclusion and exclusion criteria.

Inclusion Criteria
Peer-reviewed journals and conference publications dated between 2000–2018
The study reports the use of defensive approaches to solve cybersecurity problems
The study reports the use of offensive approaches to solve cybersecurity problems
The study presents full result to the research question
Exclusion Criteria
Papers NOT published as part of the main conference proceedings
Studies that only give general descriptions, failing to present experimental results
Studies that failed to give enough explanation of the experimental results
Studies with similar results published in different venues

2.4. Machine Learning Techniques Primer

To answer the research question about the prevalence of ML techniques, we divided the ML methods within defensive and offensive cybersecurity. In general, ML techniques can be categorized into three main groups: supervised, unsupervised, and semi-supervised learning. With supervised learning, the ML algorithms require prior knowledge to guide decisions. This includes detailed data from past security incidents with an assigned label as to whether this was a breach or not. This popular type of supervised ML method is called a classifier. For instance, the training data could include information about network packets sent during an attack, along with other properties such as originating source details. The patterns within the training data is then associated with a "threat" or "no threat" label by the ML algorithm, and the trained model can classify future unknown threats. On the other hand, unsupervised ML methods do not rely on training data or curated labels, but group threats and non-threats based on general-purpose patterns within observations. One popular unsupervised ML method is clustering, where data points with similar attributes are grouped together, such as signals for attacks as an example. One benefit of unsupervised ML is that historical data is not needed for training. On the other hand, unsupervised algorithms tend to be more general-purpose compared with supervised methods that can learn domain-specific data properties better. For example, an unsupervised method for malware detection may take longer to detect new threats until the algorithm parameters are changed by the domain experts. On the other hand, supervised ML algorithms would be able to correlate new threats as soon as new signatures are provided as training data. Semi-supervised ML is useful when the training data is insufficient for supervised ML, but the unsupervised alternative may not give the best results. In this scenario, a small curated data set of attack signals can be used to make temporary inferences about new signals in conjunction with unsupervised ML approaches.

ML techniques are evaluated using standard metrics such as true positive rate, false positive rate, accuracy, precision, recall, F1 score, false alarm rate, and confusion matrix. The basis of these metrics are true positives, false positives, true negatives, and false negatives. True positive rate is considered as the number of intrusions that were correctly detected over the total number of intrusions in the testing set. False positive rate is a representation of the number of normal requests recognized as intrusions over the total number of normal requests available in the testing set [11]. It should also be noted that False Positive Rate (FPR) and False Negative Rate (FNR) are essential metrics in various types of network systems, including Erdos–Rényi (ER) networks, Random Regular (RR) networks, and Scale-Free (SF) networks. FPR and FNR provide a robust measurement of a network's probability of being compromised by attacks or security threats being overlooked [12]. Accuracy provides the percentage of all requests correctly classified. That is, the number of requests correctly classified over the total number of requests available in the testing set expressed as a percentage [13]. Another common measure is precision, computed as the number of intrusions that were correctly classified over the total number of observed data points. Recall is another measure to compute the ratio of the number of correctly identified intrusions over the total number of intrusions. F1 score is a composite measure computed as the weighted average of the recall and precision [14]. It provides a balance through incorporation of both precision and recall. False alarm rate is considered as the proportion between the number of normal connections that are incorrectly categorized as attacks and the aggregate of normal connections [11]. The confusion matrix presents the distribution of correctly and incorrectly classified or predicted data points [14].

3. Results

Having followed the review protocol we outlined, the results of the systematic review are summarized by answering each of the three research questions we raised. Firstly, we presented results of the survey on offensive and defensive ML techniques in cybersecurity, followed by a summary of data sets used in supervised ML. Lastly, we categorized cybersecurity threats that were tackled with ML techniques in the related literature. It should be noted that there is no silver bullet classification

algorithm, and specific context requires thorough evaluation to determine the best classifiers suitable for the cybersecurity classification problem at hand.

Generally, most of the authors of the selected reviewed papers employed defensive approaches to provide solutions to the various cybersecurity issues and this can be seen based on Figure 2. Offensive approaches in cybersecurity were first employed in few of the selected reviewed papers published in 2008 and the number of papers that used these approaches comparatively increased in 2012. Despite this increase, the number papers that used defensive approaches in 2012 was about three times the number of papers leveraging offensive approaches in that same year. Furthermore, the selected reviewed papers published in 2017 recorded the highest number of studies that leveraged defensive approaches to solve cybersecurity problems. We postulated that a key reason for this discrepancy might be perceptions of Return on Investment (ROI) for offensive approaches, including unclear metrics for success. Offensive methods mitigate attacks and breaches before they occur. Hence, the effectiveness of these approaches is comparatively harder to quantify with traditional security metrics like FPR and FNR.

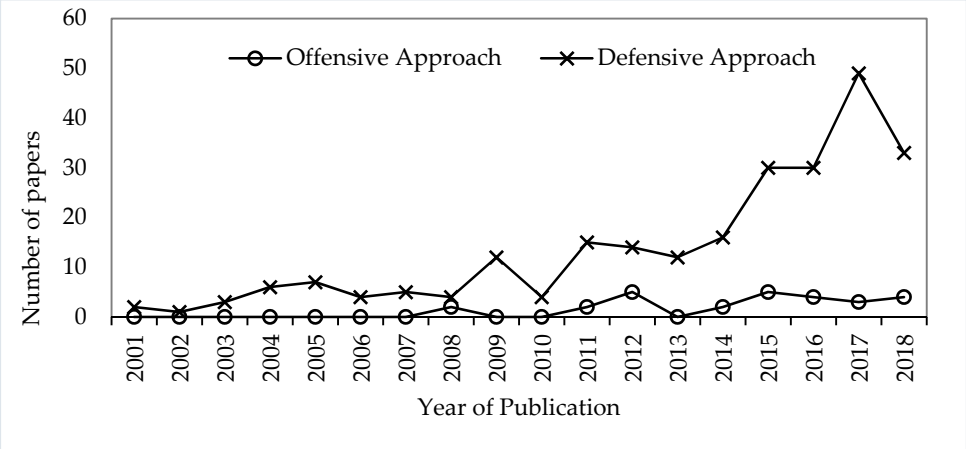

Figure 2. Trend for Machine Learning (ML) -based offensive and defensive approaches in cybersecurity.

3.1. Defensive Machine Learning Techniques

We identified seven commonly used classification techniques in the selected reviewed papers on IDS, a defensive cybersecurity strategy: Support Vector Machine (SVM), naïve Bayes, decision trees, random forests, logistic regression, neural networks, and hybrid methods. Figure 3 summarizes the range and prevalence of ML defensive strategies surfaced in literature.

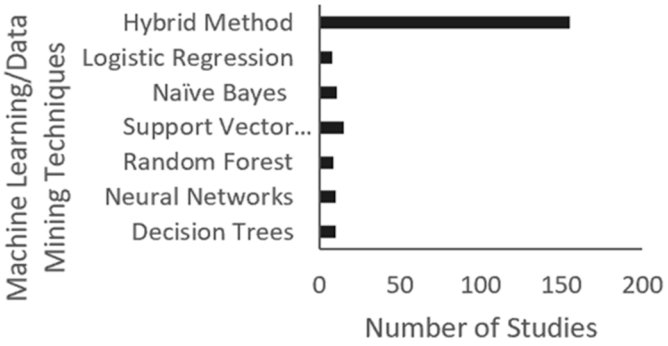

Figure 3. Machine Learning (ML) techniques used in defensive approaches.

SVM are a frequently used classification algorithm by authors of the selected reviewed papers when dealing with IDS. This classification algorithm searches for an optimum hyperplane to divide

two classes by conducting a structural risk analysis of statistical learning theory [15]. For the definition of this hyperplane, the algorithm ensures some support vectors are computed so that the maximum margin could be achieved. Some authors applied a soft margin when the data set was imperfectly linearly separable. This means that the authors were able to change the non-linear support vector machines to a linear problem using kernel functions. Many of the reviewed papers employed radial basis function (RBF). RBF has satisfactory non-linear forecasting abilities and RBF SVM has a smaller number of controllable parameters with respect to linear SVMs.

Naïve Bayes classifier is considered the simplest form of Bayesian network classifiers as all attributes are naively assumed to be unconstrained. Several authors have used this classifier in many studies because of its accuracy, performance, as well as simplicity, which can be attributed to its assumption property, which is conditionally independent. However, some authors have discovered that this classification algorithm will not have a good performance if there is an unsatisfied assumption property especially with data sets such as the KDD'99 data set which has complex attribute dependencies [16].

Decision trees are another widely used classification technique consisting of leaf nodes and decision nodes. One major evaluation factor for decision trees is classification error. This error has been defined as the misclassified cases percentage. When the class categories are more in the decision tree, there is a significant reduction in the classification accuracy [16]. Some authors considered the decision tree-based algorithms to be more advantageous than SVM as decision tree-based algorithms, especially J48, showed better weighted recall and overall accuracy. Furthermore, those authors concluded that decision tree-based algorithms provide better understanding of various classes of malicious behaviors as results were better interpreted.

Another common classification technique is random forest. Through this technique, several trees from the training data set are created. Every data set will go through the forest of trees to be classified and by averaging prediction from all the trees, the results are calculated. This classification technique has been considered to have excellent accuracy. Random forest has also been shown to help reduce false alarms and processing times [16].

Logistic regression is a probabilistic linear classifier that involves the projection of input vectors onto hyperplanes. The probability that the input is a member of a corresponding class is reflected through the distance of the input to the hyperplane. Though the logistic classifier needs extensive training time in some instances, it is an efficient classifier and has been used in cybersecurity to more effectively handle noisy data that can often be generated when trying to deal with security threats and attacks [17].

Neural networks were also used in many of the selected reviewed papers for IDS. More advanced variants of this classification technique include deep learning, with the support of several layers of connected networks. This classification algorithm is suitable for solving complicated data problems by extraction of sophisticated patterns from features with limited prior knowledge. Backpropagation is a common method used in training a neural network, and has been considered to result in local solutions or cause low training speed so a single-hidden layer feed-forward neural network, called extreme learning machines (ELMs) was proposed. The ELMs use bias and random weights for the connection of hidden neurons and input. Also, the ELMs use only a one step calculation of least squares approximation to determine the weights, thereby speeding up the learning process [18].

Finally, many current IDS studies have proposed hybrid detection techniques, combining both signature-based detection techniques and anomaly-based detection techniques. Several papers integrated classification algorithms to create effective IDS. One of the papers of interest proposed a Signature-based Anomaly Detection System (SADS) to overcome some of the drawbacks of the conventional IDS, such as false alarms, by integrating naïve Bayes and random forest classifiers [16].

It was also discovered that when random forest and NBTree algorithms, which combine naïve Bayes and decision tree classifiers, were used cooperatively based on the sum rule scheme, the detection accuracy was greater than the singular random tree algorithm's detection accuracy [19]. Malicious web

sessions have also been automatically classified to multiple vulnerability scan classes and attack classes using various multiclass supervised ML methods such as SVM, J48, and PART [20].

3.2. Offensive Machine Learning Techniques

Figure 4 shows the popular techniques used in offensive cybersecurity. Neural networks were the more commonly used technique, while unsupervised ML techniques such as association rule mining [21], frequent pattern mining [11], and clustering [22] were used in a few studies. Combinations of supervised and unsupervised methods as semi-supervised learning were also used [23].

Figure 4. Machine Learning (ML) techniques used in offensive approaches.

Many authors have conducted research on bigraph-based ML algorithms and a subset of anonymized data learning methods function as offensive filter-based methods for network defense as well as moving target defense techniques that change the view of the network from the attacker through spatio-temporal randomization. These studies contributed significantly to a paradigm shift from defensive to offensive cybersecurity. One of the selected reviewed papers of interest discussed a framework created by the authors with the purpose of predicting adversarial movement with progressing threats. The authors employed the Cloppert 12-stage intrusion-chain model [24] and used various ML and DM models for predicting threats with time series data: Auto Regressive Integrated Moving Average (ARIMA), Nonlinear Auto Regressive (NAR) neural network, NAR neural network with eXogenous input (NARX), and NAR neural network for multi-steps ahead prediction. The authors were successful in predicting adversarial movement with the proposed innovative mixed-methods approach [25]. Similarly, to solve the challenge of predicting potentially malicious actions in execution files, Recurrent Neural Network (RNN) models have been used, as RNNs have been proven to be effective in time series data processing, providing high accuracy with minimal execution time for the case of dynamic malware, with appropriate feature selection and optimally configured hyperparameters [26].

3.3. Supervised Machine Learning Data Sets

Most of the studies surveyed using supervised ML with defensive or offensive security mechanisms used contest-curated data sets, while a number of recent articles have also used real-world data. The listing of data sets for supervised ML is summarized in Figure 5.

KDDCUP'99 was the most commonly used data set in the selected reviewed papers and this was followed by DARPA'99, NSL-KDD, log files, and honeypot. In Figure 5, the list of others includes MiniChallenge2, PREDICT, ADFA-LD, UCI, USPS, CDMC2012, SEA, CSIC2010, SSNET2011, ISCX, MNIST, ISOT, HIGGS1, SUSY2, Collection, Flower, NAB, MUTAG, ENZYMES, Tiki Usenet, CIDDS-001, Netresec AB 2015, RTBTE, CAIDA'07, and CAIDA'08.

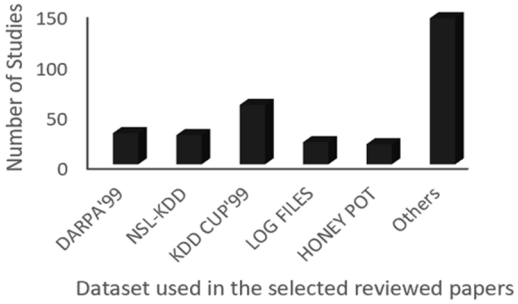

Figure 5. Data set used with supervised Machine Learning (ML) approaches.

DARPA'99 data set was commonly used in the reviewed papers when testing IDS methods as it is considered well documented and well-studied in IDS. This data set is widely used especially in articles that focused on signature-based IDS. This type of detection involves the availability of a normal network traffic and a whole set of attack with each attack having a comprehensive information such as source and destination ports, source and destination IP addresses, attack duration, attack starting time, and other relevant information [27]. Though DARPA'99 data set is considered well documented and well-studied, many of the reviewed papers combined it with other data sets when evaluating intrusion methods [16]. This was because many researchers concluded that DARPA'99 is old and not suitable for evaluating recent IDS methods. Some of the reviewed articles either used both DARPA'99 and UNSW-NB15 Data Set or used only the UNSW-NB15 Data Set which is a recent data set generated with the aim of responding to the inaccessibility of network benchmark data set challenges [27].

The NSL-KDD data set is an updated version of KDDCUP'99 and was used in many of the articles reviewed as researchers consider KDDCUP'99 to have some innate flaws. One major flaw is the large number of unwanted records as duplicates of records were found when the KDDCUP'99 was used for IDS. This flaw causes bias in evaluation results [28]. However, many recent articles employ KDDCUP'99 when evaluating the performance of anomaly-based IDS. Also, many authors consider the KDDCUP'99 to provide non-biased evaluation results especially when used for the evaluation of anomaly-based IDS performance [19]. NSL-KDD is still regarded to have some problems, such as unrealistic data rates of normal and attack data. Nevertheless, many studies employ NSL-KDD data set when evaluating IDS approaches because its test subset and train subset records are reasonable [29].

Log files were also employed in the reviewed literature. Though log files were not as commonly used in the reviewed papers as DARPA'99, KDDCUP'99, and NSL-KDD, it was still employed in a significant number of studies. Some of the studies used log files obtained from anti-virus logs, firewall servers, and IDS [30]. Web logs as real-world data sets contained various websites and relevant information about ATTACK-DATE, HOST, PARAMETERS, URL, REFER, USER-AGENT, COOKIE, IP, POST-CONTENT, as well as other details.

Various honeypot mechanisms provided as data sets were also leveraged in the surveyed research papers, which could be deployed to different networks to provide effective and practical evaluation results. Most of the reviewed papers that employed honeypots when evaluating IDS approaches made some modifications in addition to the traditional honeypots. For example, in one of the reviewed papers, the researchers deployed offensive systems that access joint botnets and malicious web servers to receive different types of commands as the traditional honeypots only receive attacks. There were also some less common real-world data sets used in some of the reviewed papers. Some of the less common real-world data sets were a real-world benchmark corpus which contained an estimate of one billion words from the Google code project and the real-world data sets from the ML database repository of the University of California, Irvine (UCI) [31].

3.4. Cyberattacks Tackled by Machine Learning

In this section, an overview of the major challenges discussed in the selected reviewed papers is presented. Many authors of the selected reviewed papers employed different ML techniques to solve some of the most common IDS challenges which have been extensively discussed in IDS and cybersecurity research. Figure 6 summarizes the cyberattacks tackled by ML techniques.

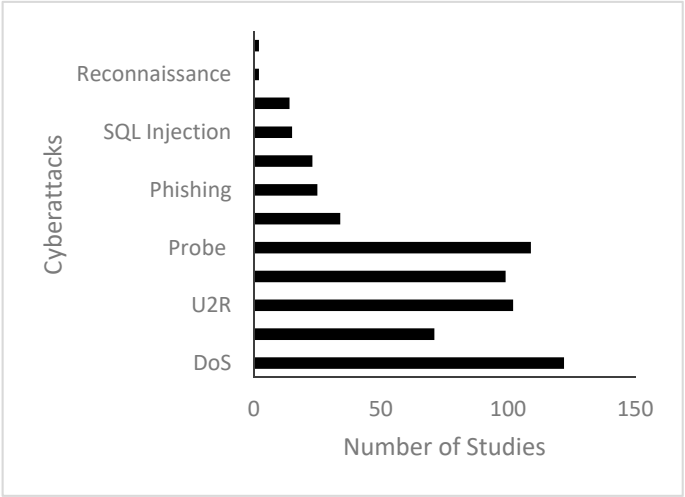

Figure 6. Cyberattacks solved by the Machine Learning (ML) techniques used in the selected reviewed papers and the number of studies that employed ML techniques to solve the different attacks.

Denial of Service (DoS), User to Root (U2R), Remote to Local (R2l), and probe attacks have been the most common categories of attacks solved with ML techniques. These attacks were found in most of the data sets (training and testing set). The DoS attack is an attack that results in the unavailability of network resources to intended users as services of a host connected to the internet becomes suspended. U2R attacks involve attackers attempting to get access of a target system without gaining official permission or approval. R2L attacks involve situations where attackers exploit vulnerabilities which could involve the guessing of passwords to take control over a remote machine. Probe attacks involve attackers examining machines to obtain relevant information [32]. Distributed Denial of Service (DDoS) attack is also a common attack solved with ML techniques. This attack has been considered the most enhanced form of DoS attacks. Its power to deploy attack vectors over the internet in a "distributed way" and generate lethal traffic through the aggregation of these forces differentiates it from other attacks [33].

Anti-Malware software products which protect legitimate users from attacks mostly use signature-based detection methods. These methods involve the extraction of unique signatures from already known malicious files and the identification of an executable file as a malicious code if there is a match between its signatures and the list containing available signatures [34]. Many studies have shown that the restriction of the signature-based detection method to recognize already known malware has made it ineffective and unreliable against new malicious codes [17].

Currently, many firms deal with botnets as it has been considered as one of the most significant cybersecurity threats. Many cybercrimes committed usually involved the use of botnets. Although, many studies have discussed different methods in which botnet could be detected and analyzed, coping with new forms of botnets has become a major challenge which has received relatively limited attention [35]. Various reviewed papers proposed new detection techniques to address botnet attacks. In some of the papers, traffic behavior analysis was used in detecting botnet activities as network traffic behavior was classified using ML. The traffic behavior analysis approach can function normally with encrypted network communication protocols as it is independent of packet payloads.

Also, information regarding network traffic can be recovered with ease from different network devices without the service availability or network performance being significantly affected. Furthermore, it has been discovered that various existing botnet detection techniques depend on the detection of botnet activities during the initial formation phase or attack phase. As a result, some of the studies proposed techniques for the detection of botnets during the initial formation phase as well as during the control and command phase.

One of the current attacks organizations face is the zero-day attack [36]. Many of the selected reviewed papers focus on anomaly detection methods to detect zero-day attacks using behavior-based data from benign programs. Some of the studies proposed a host-based anomaly detection method. In one of the studies, fuzzy logic and genetic algorithms were employed for anomaly detection. Furthermore, a significant number of the selected reviewed papers presented an enhanced or modified SVM approach to solve this challenge. Several challenges in detecting sequential data anomalies still exists even though anomaly detection techniques are employed in various studies and applied in several areas [37]. Furthermore, many of the selected reviewed papers employed ML towards cyberattacks such as malware, phishing, SQL injections, ransomware, and cross-site scripting (XSS) [38]. Authors find detecting anomalies in sequential data more complicated than detecting anomalies in static patterns and this is due to the sequential data's temporary related nature. Many authors of the selected reviewed papers propose different novel ML techniques to solve this challenge. In one of the studies, they proposed a temporal difference (TD) learning based method. The authors made some modifications to the Markov reward model which is often used to detect multi-stage cyberattacks. Furthermore, the value functions of Markov reward process were equivalent to the anomaly probabilities of sequential behaviors were included in the proposed TD learning based approach.

4. Discussion

There are a number of insights and challenges that can be identified in the use of ML with cybersecurity. These include the high dimensionality of network traffic data, class overlap between threats and legitimate data over the feature space, and general uncertainty of information. Nevertheless, ML in cybersecurity is growing exponentially, and the future points to more utilization.

4.1. High Dimensionality of Network Traffic Data

High dimensionality of network traffic data has made classification challenging as network traffic data usually comprise of many attributes and features [39]. This is mainly due to the computational complexity and resources required to process large and sparse matrices with many feature columns and observation rows. This challenge makes it difficult for researchers to train models in order to differentiate between anomalous and normal behavior [40]. As a result, many of the selected reviewed papers discussed the need for a reduction in the dimensionality of network traffic data as well as feature selection and the introduction of ML techniques when classifying such data. Some of the authors of the selected reviewed papers employed data mining techniques in a cloud-based environment, by choosing suitable attributes and features with the least relevance with regards to weight for the classification. In the majority of the reviewed studies, the standard strategy is to choose features with more desirable weights.

4.2. Class Overlap Challenge

Network intrusion detection systems face the challenge of providing satisfactory detection results due to class overlap between threat and legitimate data over feature space. This is also one of the causes behind false alarms and false positives observed in ML-based IDS. Another aspect of the class overlap challenge is the temporal shifting of network nodes from being threats to non-threats [41]. Nodes can be malicious or non-malicious at different times due to changes in performance, resource availability, or infections, disinfections, and re-infections. Various authors of the selected reviewed papers proposed different ML optimizations to solve this class overlap challenge. One of the studies introduced a wavelet

based multi-scale Hebbian learning approach to neural networks [42], and the proposed methodology was able to properly differentiate between non-linear and overlapping boundaries.

4.3. Uncertainty of Information

The increasing occurrence of cyberattacks worldwide has resulted in the misuse or loss of information assets, thereby, increasing organizations' expenses [43]. Over the years, intrusion detection systems have been used for the protection of networks and computer systems. To detect cyberattacks, most of the present intrusion detection systems depend on low-level raw network data. A current practice is to employ knowledge-based intrusion detection systems which store cyberattack related information as well as the corresponding vulnerabilities. Also, this stored information is used for guiding the process of predicting attacks. A major challenge knowledge-based IDS face is the inability to predict attacks due to the lack of contextual information or the uncertainty of information. Contextual information involves not only information about the configuration on the target systems and their vulnerabilities but any important pre-condition that must exist to achieve a successful attack. Also, contextual information includes probable semantic relationships between the targeted locations and the activities of the attackers at the time of the activities. Machine learning and probabilistic approaches have been employed in many studies to tackle common uncertainty challenges. However, many authors discovered that these approaches use models that users cannot understand but fuzzy logic approaches model uncertainty in a user-friendly form.

4.4. Future of Machine Learning in Cybersecurity

Looking ahead, some of the predictions around new cybersecurity threats involve the exploitation of ML systems by attackers and the use of these ML systems to aid assaults. Even though ML systems have been useful in automating manual activities and enhancing decision-making, they are also targets of new attacks. The fragility of some ML technologies has been predicted to become a growing concern. ML systems are also potential targets by hackers, and ML techniques can be used by attackers to enhance their attack vectors and data sniffing activities. Also, phishing and other social engineering attacks could be made better using ML, fooling targeted individuals through the creation of well-crafted audio-visuals or untraceable emails. Furthermore, realistic disinformation campaigns could be launched using ML. The generation of new threats can been made relatively easy for attackers due to the availability of attack toolkits for sale online. Another prediction made by some researchers relating to cybersecurity is increased dependence on ML for countering attacks and identifying vulnerabilities. Mobile phone users could be warned of risky actions when ML is embedded into mobile phones. Trade-offs between tracking personal information in exchange for added security is an ongoing discussion especially within research on security-based ML [44].

5. Conclusions

In the fight against malicious threats, there has been collaborative support from experts to design different cyber defense systems. Intrusion detection mechanisms monitor, track, and block viruses and other malicious cyberattacks. However, these methods are still vulnerable to attacks in applications because the design and implementation of software and networks is imperfect. Advanced methods using ML are being developed to discover previously unknown cyber intrusions and techniques towards a more dependable cybersecurity infrastructure, including both defensive and offensive approaches. This paper systematically synthesized the knowledge base in the domain of cybersecurity with ML. We also covered current challenges and future directions of ML in cybersecurity while surveying nearly two decades of research in the applications of ML to security. By employing the Systematic Literature Reviews and PRISMA model, we answered three research questions on ML techniques being used in offensive and defensive cybersecurity, data sets being used in training supervised ML models, as well as cyberattacks that have been tackled by ML. Our study was limited to literature investigation, and while algorithmic and experimental comparison of the different approaches

was beyond our scope, this would be an interesting research direction for future work in this area. Although there is no silver bullet ML algorithm to handle all possible cybersecurity vulnerabilities, threats, and attacks, our study shows impressive outcomes from ML solutions, and provides a good starting point for researchers exploring ML techniques within cybersecurity.

Author Contributions: Conceptualization, I.D.A.; methodology, I.D.A. and H.S.; validation, I.D.A. and H.L.; formal analysis, I.D.A., H.S., and H.L.; investigation, H.L.; resources, I.D.A. and H.L.; data curation, I.D.A. and H.L.; writing—original draft preparation, I.D.A.; writing—review and editing, H.S.; visualization, I.D.A. and H.S.; supervision, H.L.; project administration, H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

Appendix A

Table A1. PRISMA Checklist [9].

Section/Topic	#	Checklist Item	Reported on Page #
		TITLE	
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
		ABSTRACT	
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	1
		INTRODUCTION	
Rationale	3	Describe the rationale for the review in the context of what is already known.	1
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	2
		METHODS	
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	N/A
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	4
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	3
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	4
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	3–4

Table A1. *Cont.*

Section/Topic	#	Checklist Item	Reported on Page #
		TITLE	
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	N/A
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	3
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	3
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	N/A
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis.	2
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	3
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	N/A
		RESULTS	
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	3
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	21–28
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	N/A
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	5–11
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	11–12
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	N/A
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	N/A
		DISCUSSION	
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	11
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	12
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	12
		FUNDING	
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	13

Appendix B

Table A2. Enumeration of reviewed studies (Retrieved 14 February 2019).

ID	Reference	Publication Year	Title	Citations
A1	[45]	2007	D-SCIDS: Distributed Soft Computing Intrusion Detection System	207
A2	[46]	2016	Hybrid Intrusion Detection Method to Increase Anomaly Detection by Using Data Mining Techniques	1
A3	[33]	2016	Data Randomization and Cluster-Based Partitioning for Botnet Intrusion Detection	35
A4	[40]	2018	An Evaluation of the Performance of Restricted Boltzmann Machines as a Model for Anomaly Network Intrusion Detection	0
A5	[43]	2017	Contextual Information Fusion for Intrusion Detection: A Survey and Taxonomy	7
A6	[47]	2011	A New Approach Based on Honeybee to Improve Intrusion Detection System Using Neural Network and Bees Algorithm	6
A7	[21]	2017	Network Intrusion Detection for Cyber Security using Unsupervised Deep Learning Approaches	2
A8	[48]	2016	A Comparative Study of Different Fuzzy Classifiers for Cloud Intrusion Detection Systems Alerts	1
A9	[17]	2016	Toward an Online Anomaly Intrusion Detection System Based on Deep Learning	17
A10	[49]	2018	Adversarial Anomaly Detection Using Centroid-Based Clustering	0
A11	[50]	2017	Fuzziness Based Semi-Supervised Learning Approach for Intrusion Detection System	152
A12	[23]	2017	Unsupervised Labeling for Supervised Anomaly Detection in Enterprise and Cloud Networks	6
A13	[51]	2014	Supervised Learning to Detect DDoS Attacks	16
A14	[52]	2016	An Effective Intrusion Detection Framework based on MCLP/SVM Optimized by Time-Varying Chaos Particle Swarm Optimization	46
A15	[53]	2018	Performance Evaluation of Intrusion Detection based on Machine Learning using Apache Spark	2
A16	[54]	2011	TVi: A Visual Querying System for Network Monitoring and Anomaly Detection	32
A17	[27]	2018	An Auto-Learning Approach for Network Intrusion Detection	0
A18	[55]	2012	Towards a Multiagent-Based Distributed Intrusion Detection System Using Data Mining Approaches	23
A19	[56]	2015	Anomaly Detection from Log Files Using Data Mining Techniques	8
A20	[22]	2007	Adaptive Real-time Anomaly Detection with Incremental Clustering	59
A21	[57]	2011	An Evolutionary Multi-Agent Approach to Anomaly Detection and Cyber Defense	5
A22	[58]	2016	A Comparative Analysis of SVM and its Stacking with other Classification algorithm for Intrusion Detection	17
A23	[59]	2016	Multilayer Hybrid Strategy for Phishing Email Zero-Day Filtering	2

Table A2. *Cont.*

ID	Reference	Publication Year	Title	Citations
A24	[24]	2009	Security Intelligence	13
A25	[39]	2017	A Heuristic Attack Detection Approach using the 'Least Weighted' Attributes for Cyber Security Data	0
A26	[60]	2015	A Study on Intrusion Detection using Neural Networks Trained with Evolutionary Algorithms	22
A27	[61]	2011	Data Preprocessing for Anomaly Based Network Intrusion Detection: A Review	181
A28	[62]	2004	Probabilistic Inference Strategy in Distributed Intrusion Detection Systems	1
A29	[11]	2014	Network Anomaly Detection Approach based on Frequent Pattern Mining Technique	6
A30	[63]	2015	Multilayered Database Intrusion Detection System for Detecting Malicious Behaviors in Big Data Transaction	6
A31	[64]	2017	A Multi-Objective Evolutionary Fuzzy System to Obtain a Broad and Accurate Set of Solutions in Intrusion Detection Systems	0
A32	[65]	2015	Survey of Uses of Evolutionary Computation Algorithms and Swarm Intelligence for Network Intrusion Detection	5
A33	[15]	2015	An Improved Bat Algorithm Driven by Support Vector Machines for Intrusion Detection	8
A34	[66]	2012	Securing Advanced Metering Infrastructure Using Intrusion Detection System with Data Stream Mining	55
A35	[67]	2014	Mining Network Data for Intrusion Detection through Combining SVMs with Ant Colony Networks	147
A36	[68]	2005	GP Ensemble for Distributed Intrusion Detection Systems	65
A37	[36]	2015	MARK-ELM: Application of a Novel Multiple Kernel Learning Framework for Improving the Robustness of Network Intrusion Detection	45
A38	[20]	2012	Using Multiclass Machine Learning Methods to Classify Malicious Behaviors Aimed at Web Systems	13
A39	[69]	2011	Intrusion Detection using Neural Based Hybrid Classification Methods	76
A40	[70]	2017	Detecting Anomalous Behavior in Cloud Servers by Nested Arc Hidden SEMI-Markov Model with State Summarization	3
A41	[29]	2018	A Hybrid Intrusion Detection System based on ABC-AFS Algorithm for Misuse and Anomaly Detection	4
A42	[71]	2016	Intrusion Detection System Based on Cost Based Support Vector Machine	0
A43	[72]	2017	AHEad: Privacy-Preserving Online Behavioral Advertising using Homomorphic Encryption	2
A44	[73]	2016	Deep4MalDroid: A Deep Learning Framework for Android Malware Detection Based on Linux Kernel	22
A45	[74]	2013	Network Anomaly Classification by Support Vector Classifiers Ensemble and Non-linear Projection Techniques	19
A46	[75]	2014	Advocating the Use of Fuzzy Reasoning Spiking Neural P system in Intrusion Detection	3

Table A2. *Cont.*

ID	Reference	Publication Year	Title	Citations
A47	[76]	2013	An Agent-Based Approach for Building an Intrusion Detection System	6
A48	[77]	2012	A-GHSOM: An Adaptive Growing Hierarchical Self Organizing Map for Network Anomaly Detection	28
A49	[78]	2018	Efficient Privacy-Preserving Machine Learning in Hierarchical Distributed System	0
A50	[79]	2018	Rst-Rf: A Hybrid Model based on Rough Set Theory and Random Forest for Network Intrusion Detection	0
A51	[80]	2017	A LogitBoost-Based Algorithm for Detecting Known and Unknown Web Attacks	3
A52	[19]	2016	An Effective Combining Classifier Approach using Tree Algorithms for Network Intrusion Detection	38
A53	[81]	2015	On Copulas-Based Classification Method for Intrusion Detection	0
A54	[82]	2012	An Incremental Semi Rule-Based Learning Model for Cybersecurity in Cyberinfrastructures	0
A55	[83]	2012	A Network Intrusion Detection System based on a Hidden Naïve Bayes Multiclass Classifier	176
A56	[84]	2016	Data Analytics on Network Traffic Flows for Botnet Behavior Detection	10
A57	[85]	2016	MVPSys: Toward Practical Multi-View based False Alarm Reduction System in Network Intrusion Detection	9
A58	[13]	2017	Anomaly-Based Web Attack Detection: A Deep Learning Approach	0
A59	[86]	2016	Intrusion Detection Based on IDBM	3
A60	[34]	2017	A Hybrid Technique using Binary Particle Swarm Optimization and Decision Tree Pruning for Network Intrusion Detection	1
A61	[87]	2015	Study on Implementation of Machine Learning Methods Combination for Improving Attacks Detection Accuracy on Intrusion Detection System (IDS)	6
A62	[88]	2015	Host-Based Intrusion Detection System for Secure Human-Centric Computing	3
A63	[89]	2003	A Comparative Study of Techniques for Intrusion Detection	80
A64	[90]	2003	Intrusion Detection Using Ensemble of Soft Computing Paradigms	104
A65	[91]	2005	Model Selection for Kernel Based Intrusion Detection Systems	24
A66	[92]	2004	Intrusion Detection Systems Using Adaptive Regression Spines	97
A67	[93]	2011	An Efficient Local Region and Clustering-Based Ensemble System for Intrusion Detection	12
A68	[94]	2015	Probabilistic Models-Based Intrusion Detection using Sequence Characteristics in Control System Communication	4
A69	[95]	2009	An Empirical Approach to Modeling Uncertainty in Intrusion Analysis	27
A70	[32]	2015	Two-Tier Network Anomaly Detection Model: A Machine Learning Approach	23

Table A2. *Cont.*

ID	Reference	Publication Year	Title	Citations
A71	[96]	2016	Multilayer Perceptron Algorithms for Cyberattack Detection	1
A72	[31]	2018	Partition-Aware Scalable Outlier Detection Using Unsupervised Learning	0
A73	[97]	2012	A Novel Multi-Threaded K-Means Clustering Approach for Intrusion Detection	10
A74	[98]	2017	Using Google Analytics to Support Cybersecurity Forensics	0
A75	[26]	2018	Early-Stage Malware Prediction using Recurrent Neural Networks	5
A76	[99]	2018	Adaptive and Online Network Intrusion Detection System using Clustering and Extreme Learning Machines	4
A77	[100]	2010	Optimal Bayesian Network Design for Efficient Intrusion Detection	4
A78	[101]	2018	A Bi-Objective Hyper-Heuristic Support Vector Machines for Big Data Cyber-Security	3
A79	[14]	2015	Web Service Intrusion Detection Using a Probabilistic Framework	1
A80	[102]	2016	Computational Intelligence in Intrusion Detection System for Snort Log using Hadoop	0
A81	[103]	2016	Automated Intelligent Multinomial Classification of Malware Species using Dynamic Behavioral Analysis	3
A82	[104]	2017	Demystifying Numenta Anomaly Benchmark	3
A83	[105]	2015	An Intrusion Detection System using Network Traffic Profiling and Online Sequential Extreme Learning Machine	67
A84	[106]	2013	An Ensemble Approach for Cyber Attack Detection System: A Generic Framework	22
A85	[107]	2013	Toward a More Practical Unsupervised Anomaly Detection System	73
A86	[108]	2015	Intrusion Detection System using Bagging Ensemble Selection	2
A87	[109]	2016	A Cross-Domain Comparable Measurement Framework to Quantify Intrusion Detection Effectiveness	0
A88	[110]	2013	An Approach to the Correlation of Security Events based on Machine Learning Techniques	14
A89	[111]	2017	A Learning-Based Hybrid Framework for Detection and Defense of DDoS Attacks	0
A90	[112]	2015	Advanced Temporal-Difference Learning for Intrusion Detection	3
A91	[113]	2015	A New Privacy-Preserving Proximal Support Vector Machine for Classification of Vertically Partitioned Data	23
A92	[30]	2017	Analyst Intuition Inspired High Velocity Big Data Analysis using PCA Ranked Fuzzy K-Means Clustering with Multi-Layer Perceptron (MLP) to Obviate Cyber Security Risk	0
A93	[44]	2018	Cyber Security Predictions: 2019 and Beyond	0
A94	[114]	2014	A Fuzzy Intrusion Detection System based on Categorization of Attacks	0
A95	[115]	2016	Local Outlier Factor and Stronger One Class Classifier Based Hierarchical Model for Detection of Attacks in Network Intrusion Detection Data set	3
A96	[116]	2018	Adaptive Artificial Immune Networks for Mitigating DoS Flooding Attacks	8

Table A2. *Cont.*

ID	Reference	Publication Year	Title	Citations
A97	[117]	2017	Applying Convolutional Neural Network for Network Intrusion Detection	10
A98	[118]	2017	Evaluating Effectiveness of Shallow and Deep Networks to Intrusion Detection System	4
A99	[119]	2013	Deconstructing the Assessment of Anomaly-based Intrusion Detectors	6
A100	[18]	2018	Network Intrusion Detection using Equality Constrained-Optimization-Based Extreme Learning Machines	2
A101	[120]	2018	Deep Learning-Based Intrusion Detection With Adversaries	6
A102	[121]	2014	An Evasion and Counter-Evasion Study in Malicious Websites Detection	24
A103	[37]	2010	Sequential Anomaly Detection based on Temporal-Difference Learning: Principles, Models and Case Studies	49
A104	[122]	2017	Continuous Implicit Authentication for Mobile Devices based on Adaptive Neuro-Fuzzy Inference System	2
A105	[16]	2014	Signature-Based Anomaly Intrusion Detection using Integrated Data Mining Classifiers	11
A106	[123]	2015	Privacy-Preserving Association Rule Mining in Cloud Computing	26
A107	[124]	2017	A Binary-Classification Method Based on Dictionary Learning and ADMM for Network Intrusion Detection	1
A108	[125]	2011	An Effective Network-Based Intrusion Detection using Conserved Self Pattern Recognition Algorithm Augmented with Near-Deterministic Detector Generation	10
A119	[126]	2008	ULISSE, A Network Intrusion Detection System	12
A110	[127]	2016	Causality Reasoning about Network Events for Detecting Stealthy Malware Activities	26
A111	[128]	2015	A Novel Anomaly Detection Approach for Mitigating Web-Based Attacks Against Clouds	4
A112	[129]	2011	Artificial Immune System-Based Intrusion Detection in a Distributed Hierarchical Network Architecture of Smart Grid	30
A113	[35]	2013	Botnet Detection based on Traffic Behavior Analysis and Flow Intervals	177
A114	[130]	2017	Network Intrusion Detection using Word Embeddings	1
A115	[131]	2015	Secure Multi-party Computation Based Privacy Preserving Extreme Learning Machine Algorithm Over Vertically Distributed Data	5
A116	[132]	2011	On the Design and Analysis of the Privacy-Preserving SVM Classifier	146
A117	[25]	2018	Predicting Adversarial Cyber-Intrusion Stages Using Autoregressive Neural Networks	0
A118	[28]	2010	Discriminative Multinomial Naïve Bayes for Network Intrusion Detection	71
A119	[42]	2017	Multiscale Hebbian Neural Network for Cyber Threat Detection	2
A120	[133]	2011	Detecting P2P Botnets through Network Behavior Analysis and Machine Learning	174

References

1. Dua, S.; Du, X. *Data Mining and Machine Learning in Cybersecurity*; Auerbach Publications: Boca Raton, FL, USA, 2016.
2. Triplett, C. Security is Only as Strong as the Weakest Link. *Infosecurity Magazine*, 21 November 2019. Available online: http://www.infosecurity-magazine.com/opinions/strong-weakest-link (accessed on 18 August 2020).
3. IEEE. *Artificial Intelligence and Machine Learning Applied to Cybersecurity*; IEEE: New York, NY, USA, 2017.
4. Nimon, H.I. *Offensive and Defensive Security: Concepts, Planning, Operations, and Management*; Xlibris Corporation: Bloomington, IN, USA, 2013.
5. Sollaci, L.B.; Pereira, M.G. The Introduction, Methods, Results, and Discussion (IMRAD) Structure: A Fifty-Year Survey. *J. Med. Libr. Assoc.* **2004**, *92*, 364–371.
6. Dilek, S.; Çakır, H.; Aydın, M. Applications of Artificial Intelligence Techniques to Combating Cyber Crimes: A Review. *arXiv* **2015**, arXiv:1502.03552.
7. Li, J.-H. Cyber Security Meets Artificial Intelligence: A Survey. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, h1462–1474. [CrossRef]
8. Charters, S.; Kitchenham, B.A. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; EBSE Technical Report; University of Durham: Durha, UK, 2007.
9. Liberati, A.; Altman, D.G.; Tetzlaff, J.; Mulrow, C.; Gøtzsche, P.C.; Ioannidis, J.P.A.; Clarke, M.; Devereaux, P.J.; Kleijnen, J.; Moher, D. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Med.* **2009**, *6*, e1000100. [CrossRef] [PubMed]
10. Shang, Y. Subgraph Robustness of Complex Networks under Attacks. *IEEE Trans. Syst. Man Cybern. Syst.* **2017**, *49*, 821–832. [CrossRef]
11. Dominic, D.D.; Said, A.M. Network Anomaly Detection Approach based on Frequent Pattern Mining Technique. In Proceedings of the International Conference on Computational Science and Technology (ICCST), Kota Kinabalu, Malaysia, 27–28 August 2014; pp. 1–6.
12. Shang, Y. False Positive and False Negative Effects on Network Attacks. *J. Stat. Phys.* **2018**, *170*, 141–164. [CrossRef]
13. Liang, J.; Zhao, W.; Ye, W. Anomaly-Based Web Attack Detection: A Deep Learning Approach. In Proceedings of the International Conference on Network, Communication and Computing (ICNCC), Silicon Valley, CA, USA, 26–29 January 2017; pp. 80–85.
14. Sallay, H.; Bourouis, S.; Bouguila, N. Web Service Intrusion Detection Using a Probabilistic Framework. In *Progress in Systems Engineering Advances in Intelligent Systems and Computing*; Springer International Publishing: Cham, Switzerland, 2015; pp. 161–166.
15. Enache, A.-C.; Sgârciu, V. An Improved Bat Algorithm Driven by Support Vector Machines for Intrusion Detection. In Proceedings of the Advances in Intelligent Systems and Computing International Joint Conference, Burgos, Spain, 15–17 June 2015; pp. 41–51.
16. Yassin, W.; Udzir, N.I.; Abdullah, A.; Abdullah, M.T.; Zulzalil, H.; Muda, Z. Signature-Based Anomaly Intrusion Detection using Integrated Data Mining Classifiers. In Proceedings of the International Symposium on Biometrics and Security Technologies (ISBAST), Kuala Lumpur, Malaysia, 26–27 August 2014; pp. 232–237.
17. Alrawashdeh, K.; Purdy, C. Toward an Online Anomaly Intrusion Detection System Based on Deep Learning. In Proceedings of the IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2016; pp. 195–200.
18. Wang, C.-R.; Xu, R.-F.; Lee, S.-J.; Lee, C.-H. Network Intrusion Detection using Equality Constrained Optimization-Based Extreme Learning Machines. *Knowl. Based Syst.* **2018**, *147*, 68–80. [CrossRef]
19. Kevric, J.; Jukic, S.; Subasi, A. An Effective Combining Classifier Approach using Tree Algorithms for Network Intrusion Detection. *Neural Comput. Appl.* **2016**, *28*, 1051–1058. [CrossRef]
20. Goseva-Popstojanova, K.; Anastasovski, G.; Pantev, R. Using Multiclass Machine Learning Methods to Classify Malicious Behaviors Aimed at Web Systems. In Proceedings of the IEEE International Symposium on Software Reliability Engineering, Dallas, TX, USA, 27–30 November 2012; pp. 81–90.

21. Alom, M.Z.; Taha, T.M. Network Intrusion Detection for Cyber Security using Unsupervised Deep Learning Approaches. In Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON), Dayton, OH, USA, 15–19 July 2017; pp. 63–69.
22. Burbeck, K.; Nadjm-Tehrani, S. Adaptive Real-Time Anomaly Detection with Incremental Clustering. *Inf. Secur. Tech. Rep.* **2007**, *12*, 56–67. [CrossRef]
23. Baek, S.; Kwon, D.; Kim, J.; Suh, S.C.; Kim, H.; Kim, I. Unsupervised Labeling for Supervised Anomaly Detection in Enterprise and Cloud Networks. In Proceedings of the IEEE International Conference on Cyber Security and Cloud Computing (CSCloud), New York, NY, USA, 26–28 June 2017; pp. 205–210.
24. Cloppert, M. Security Intelligence: Attacking the Cyber Kill Chain; SANS Computer Forensics. 2009. Available online: http://www.sans.org/blog/security-intelligence-attacking-the-cyber-kill-chain/ (accessed on 12 December 2019).
25. Rege, A.; Obradovic, Z.; Asadi, N.; Parker, E.; Pandit, R.; Masceri, N.; Singer, B. Predicting Adversarial Cyber-Intrusion Stages Using Autoregressive Neural Networks. *Ieee Intell. Syst.* **2018**, *33*, 29–39. [CrossRef]
26. Rhode, M.; Burnap, P.; Jones, K. Early-Stage Malware Prediction using Recurrent Neural Networks. *Comput. Secur.* **2018**, *77*, 578–594. [CrossRef]
27. Boulaiche, A.; Adi, K. An Auto-Learning Approach for Network Intrusion Detection. *Telecommun. Syst.* **2018**, *68*, 277–294. [CrossRef]
28. Panda, M.; Abraham, A.; Patra, M.R. Discriminative Multinomial Naïve Bayes for Network Intrusion Detection. In Proceedings of the International Conference on Information Assurance and Security, Atlanta, GA, USA, 23–25 August 2010; pp. 5–10.
29. 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]
30. Teoh, T.T.; Zhang, Y.; Nguwi, Y.Y.; Elovici, Y.; Ng, W.L. Analyst Intuition Inspired High Velocity Big Data Analysis using PCA Ranked Fuzzy k-Means Clustering with Multi-Layer Perceptron (MLP) to Obviate Cyber Security Risk. In Proceedings of the International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Guilin, China, 29–31 July 2017; pp. 1790–1793.
31. Parveen, P.; Lee, M.; Henslee, A.; Dugan, M.; Ford, B. Partition-Aware Scalable Outlier Detection Using Unsupervised Learning. In Proceedings of the IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, USA, 7–9 July 2018; pp. 186–192.
32. Pajouh, H.H.; Dastghaibyfard, G.; Hashemi, S. Two-Tier Network Anomaly Detection Model: A Machine Learning Approach. *J. Intell. Inf. Syst.* **2015**, *48*, 61–74. [CrossRef]
33. Al-Jarrah, O.Y.; Alhussein, O.; Yoo, P.D.; Muhaidat, S.; Taha, K.; Kim, K. Data Randomization and Cluster-Based Partitioning for Botnet Intrusion Detection. *IEEE Trans. Cybern.* **2016**, *46*, 1796–1806. [CrossRef]
34. Malik, A.J.; Khan, F.A. A Hybrid Technique using Binary Particle Swarm Optimization and Decision Tree Pruning for Network Intrusion Detection. *Clust. Comput.* **2017**, *21*, 667–680. [CrossRef]
35. Zhao, D.; Traore, I.; Sayed, B.; Lu, W.; Saad, S.; Ghorbani, A.; Garant, D. Botnet Detection based on Traffic Behavior Analysis and Flow Intervals. *Comput. Secur.* **2013**, *39*, 2–16. [CrossRef]
36. Fossaceca, J.M.; Mazzuchi, T.A.; Sarkani, S. MARK-ELM: Application of a Novel Multiple Kernel Learning Framework for Improving the Robustness of Network Intrusion Detection. *Expert Syst. Appl.* **2015**, *42*, 4062–4080. [CrossRef]
37. Xu, X. Sequential Anomaly Detection based on Temporal-Difference Learning: Principles, Models and Case Studies. *Appl. Soft Comput.* **2010**, *10*, 859–867. [CrossRef]
38. Kaur, G.; Malik, Y.; Samuel, H.; Jaafar, F. Detecting Blind Cross-Site Scripting Attacks Using Machine Learning. In Proceedings of the International Conference on Signal Processing and Machine Learning (SPML), Shanghai, China, 28–30 November 2018; pp. 22–25.
39. Dali, L.; Mivule, K.; El-Sayed, H. A Heuristic Attack Detection Approach using the 'Least Weighted' Attributes for Cyber Security Data. In Proceedings of the Intelligent Systems Conference (IntelliSys), London, UK, 7–8 September 2017; pp. 1067–1073.
40. Aldwairi, T.; Perera, D.; Novotny, M.A. An Evaluation of the Performance of Restricted Boltzmann Machines as a Model for Anomaly Network Intrusion Detection. *Comput. Netw.* **2018**, *144*, 111–119. [CrossRef]
41. Shang, Y. Hybrid Consensus for Averager–Copier–Voter Networks with Non-Rational Agents. *Chaos Solitons Fractals* **2018**, *110*, 244–251. [CrossRef]

42. Siddiqui, S.; Khan, M.S.; Ferens, K. Multiscale Hebbian Neural Network for Cyber Threat Detection. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 1427–1434.
43. Aleroud, A.; Karabatis, G. Contextual Information Fusion for Intrusion Detection: A Survey and Taxonomy. *Knowl. Inf. Syst.* **2017**, *52*, 563–619. [CrossRef]
44. Thompson, H.; Trilling, S. Cyber Security Predictions: 2019 and Beyond. Symantec, 28 November 2018. Available online: https://www.symantec.com/blogs/feature-stories/cyber-security-predictions-2019-and-beyond (accessed on 12 December 2019).
45. Abraham, A.; Jain, R.; Thomas, J.; Han, S.Y. D-SCIDS: Distributed Soft Computing Intrusion Detection System. *J. Netw. Comput. Appl.* **2007**, *30*, 81–98. [CrossRef]
46. Ahmad, B.; Jian, W.; Hassan, B.; Rehmatullah, S. Hybrid Intrusion Detection Method to Increase Anomaly Detection by Using Data Mining Techniques. *Int. J. Database Theory Appl.* **2016**, *9*, 231–240. [CrossRef]
47. Ali, G.A.; Jantan, A. A New Approach Based on Honeybee to Improve Intrusion Detection System Using Neural Network and Bees Algorithm. In *Software Engineering and Computer Systems Communications in Computer and Information Science*, 2011 ed.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 777–792.
48. Alqahtani, S.M.; John, R.A. Comparative Study of Different Fuzzy Classifiers for Cloud Intrusion Detection Systems Alerts. In Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, 6–9 December 2016; pp. 1–9.
49. Anindya, I.C.; Kantarcioglu, M. Adversarial Anomaly Detection Using Centroid-Based Clustering. In Proceedings of the IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, USA, 7–9 July 2018; pp. 1–8.
50. 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]
51. Balkanli, E.; Alves, J.; Zincir-Heywood, A.N. Supervised Learning to Detect DDoS Attacks. In Proceedings of the IEEE Symposium on Computational Intelligence in Cyber Security (CICS), Orlando, FL, USA, 9–12 December 2014; pp. 1–8.
52. Bamakan, S.M.H.; Wang, H.; Yingjie, T.; Shi, Y. An Effective Intrusion Detection Framework based on MCLP/SVM Optimized by Time-Varying Chaos Particle Swarm Optimization. *Neurocomputing* **2016**, *199*, 90–102. [CrossRef]
53. Belouch, M.; el Hadaj, S.; Idhammad, M. Performance Evaluation of Intrusion Detection based on Machine Learning using Apache Spark. *Procedia Comput. Sci.* **2018**, *127*, 1–6. [CrossRef]
54. Boschetti, A.; Salgarelli, L.; Muelder, C.; Ma, K.-L. TVi: A Visual Querying System for Network Monitoring and Anomaly Detection. In Proceedings of the International Symposium on Visualization for Cyber Security (VizSec), Pittsburgh, PA, USA, 20 July 2011; Association for Computing Machinery: New York, NY, USA, 2011; pp. 1–10.
55. Brahmi, I.; Yahia, S.B.; Aouadi, H.; Poncelet, P. Towards a Multiagent-Based Distributed Intrusion Detection System Using Data Mining Approaches. In *Lecture Notes in Computer Science Agents and Data Mining Interaction*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 173–194.
56. Breier, J.; Branišová, J. Anomaly Detection from Log Files Using Data Mining Techniques. In *Lecture Notes in Electrical Engineering Information Science and Applications*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 449–457.
57. Carvalho, M.; Perez, C. An Evolutionary Multi-Agent Approach to Anomaly Detection and Cyber Defense. In Proceedings of the Workshop on Cyber Security and Information Intelligence Research (CSIIRW), Oak Ridge, TN, USA, 12–14 October 2011; p. 1.
58. Chand, N.; Mishra, P.; Krishna, C.R.; Pilli, E.S.; Govil, M.C. A Comparative Analysis of SVM and its Stacking with Other Classification Algorithm for Intrusion Detection. In Proceedings of the International Conference on Advances in Computing, Communication, & Automation (ICACCA), Dehradun, India, 30 September–1 October 2016; pp. 1–6.
59. Chowdhury, M.U.; Abawajy, J.; Kelarev, A.; Hochin, T. Multilayer Hybrid Strategy for Phishing Email Zero-Day Filtering. *Concurr. Comput. Pract. Exp.* **2016**, *29*, 23. [CrossRef]
60. Dash, T. A Study on Intrusion Detection using Neural Networks Trained with Evolutionary Algorithms. *Soft Comput.* **2015**, *21*, 2687–2700. [CrossRef]

61. Davis, J.J.; Clark, A.J. Data Preprocessing for Anomaly Based Network Intrusion Detection: A Review. *Comput. Secur.* **2011**, *30*, 353–375. [CrossRef]
62. Ding, J.; Xu, S.; Krämer, B.; Bai, Y.; Chen, H.; Zhang, J. Probabilistic Inference Strategy in Distributed Intrusion Detection Systems. In *Parallel and Distributed Processing and Applications Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 835–844.
63. Doroudian, M.; Arastouie, N.; Talebi, M.; Ghanbarian, A.R. Multilayered Database Intrusion Detection System for Detecting Malicious Behaviors in Big Data Transaction. In Proceedings of the International Conference on Information Security and Cyber Forensics (InfoSec), Cape Town, South Africa, 15–17 November 2015; pp. 105–110.
64. Elhag, S.; Fernández, A.; Altalhi, A.; Alshomrani, S.; Herrera, F. A Multi-Objective Evolutionary Fuzzy System to Obtain a Broad and Accurate Set of Solutions in Intrusion Detection Systems. *Soft Comput.* **2017**, *23*, 1321–1336. [CrossRef]
65. Elsayed, S.; Sarker, R.; Essam, D. Survey of Uses of Evolutionary Computation Algorithms and Swarm Intelligence for Network Intrusion Detection. *Int. J. Comput. Intell. Appl.* **2015**, *14*, 1550025. [CrossRef]
66. Faisal, M.A.; Aung, Z.; Williams, J.R.; Sanchez, A. Securing Advanced Metering Infrastructure Using Intrusion Detection System with Data Stream Mining. In *Intelligence and Security Informatics Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 96–111.
67. Feng, W.; Zhang, Q.; Hu, G.; Huang, J.X. Mining Network Data for Intrusion Detection through Combining SVMs with Ant Colony Networks. *Future Gener. Comput. Syst.* **2014**, *37*, 127–140. [CrossRef]
68. Folino, G.; Pizzuti, C.; Spezzano, G. GP Ensemble for Distributed Intrusion Detection Systems. In *Pattern Recognition and Data Mining Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 54–62.
69. Govindarajan, M.; Chandrasekaran, R.M. Intrusion Detection using Neural Based Hybrid Classification Methods. *Comput. Netw.* **2011**, *55*, 1662–1671. [CrossRef]
70. Haider, W.; Hu, J.; Xie, Y.; Yu, X.; Wu, Q. Detecting Anomalous Behavior in Cloud Servers by Nested Arc Hidden SEMI-Markov Model with State Summarization. *IEEE Trans. Big Data* **2017**, *5*, 305–316. [CrossRef]
71. Hassan, M.R. Intrusion Detection System Based on Cost Based Support Vector Machine. In *Recent Advances in Information and Communication Technology 2016 Advances in Intelligent Systems and Computing*; Springer International Publishing: Cham, Switzerland, 2016; pp. 105–115.
72. Helsloot, L.J.; Tillem, G.; Erkin, Z. AHEad: Privacy-Preserving Online Behavioural Advertising using Homomorphic Encryption. In Proceedings of the IEEE Workshop on Information Forensics and Security (WIFS), Rennes, France, 4–7 December 2017; pp. 1–6.
73. Hou, S.; Saas, A.; Chen, L.; Ye, Y. Deep4MalDroid: A Deep Learning Framework for Android Malware Detection Based on Linux Kernel System Call Graphs. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence Workshops (WIW), Omaha, NE, USA, 13–16 October 2016; pp. 104–111.
74. Hoz, E.d.l.; Ortiz, A.; Ortega, J.; De la Hoz, E. Network Anomaly Classification by Support Vector Classifiers Ensemble and Non-linear Projection Techniques. In *Lecture Notes in Computer Science Hybrid Artificial Intelligent Systems*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 103–111.
75. Idowu, R.K.; Chandren, R.; Othman, Z.A. Advocating the Use of Fuzzy Reasoning Spiking Neural P System in Intrusion Detection. In Proceedings of the Asian Conference on Membrane Computing (ACMC), Coimbatore, India, 20–22 August 2014; pp. 1–5.
76. Ioniță, I.; Ioniță, L. An Agent-Based Approach for Building an Intrusion Detection System. In Proceedings of the International Conference on Networking in Education and Research (RoEduNet), Iasi, Romania, 26–28 September 2013; pp. 1–6.
77. Ippoliti, D.; Zhou, X. A-GHSOM: An Adaptive Growing Hierarchical Self Organizing Map for Network Anomaly Detection. *J. Parallel Distrib. Comput.* **2012**, *72*, 1576–1590. [CrossRef]
78. Jia, Q.; Guo, L.; Fang, Y.; Wang, G. Efficient Privacy-Preserving Machine Learning in Hierarchical Distributed System. *IEEE Trans. Netw. Scie. Eng.* **2018**, *6*, 599–612. [CrossRef]
79. Jiang, J.; Wang, Q.; Shi, Z.; Lv, B.; Qi, B. Rst-Rf: A Hybrid Model based on Rough Set Theory and Random Forest for Network Intrusion Detection. In Proceedings of the International Conference on Cryptography, Security and Privacy (ICCSP), Guiyang, China, 16–19 March 2018; pp. 77–81.
80. Kamarudin, M.H.; Maple, C.; Watson, T.; Safa, N.S. A LogitBoost-Based Algorithm for Detecting Known and Unknown Web Attacks. *IEEE Access* **2017**, *5*, 26190–26200. [CrossRef]

81. Khobzaoui, A.; Mesfioui, M.; Yousfate, A.; Bensaber, B.A. On Copulas-Based Classification Method for Intrusion Detection. In *IFIP Advances in Information and Communication Technology Computer Science and Its Applications*; Springer International Publishing: Cham, Switzerland, 2015; pp. 394–405.
82. Kianmehr, K. An Incremental Semi Rule-Based Learning Model for Cybersecurity in Cyberinfrastructures. In Proceedings of the International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), Bangkok, Thailand, 27–31 May 2012; pp. 123–128.
83. Koc, L.; Mazzuchi, T.A.; Sarkani, S. A Network Intrusion Detection System based on a Hidden Naïve Bayes Multiclass Classifier. *Expert Syst. Appl.* **2012**, *39*, 13492–13500. [CrossRef]
84. Le, D.C.; Zincir-Heywood, N.; Heywood, M.I. Data Analytics on Network Traffic Flows for Botnet Behaviour Detection. In Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI), Athens, Greece, 6–9 December 2016; pp. 1–7.
85. Li, W.; Meng, W.; Luo, X.; Kwok, L.F. MVPSys: Toward Practical Multi-View Based False Alarm Reduction System in Network Intrusion Detection. *Comput. Secur.* **2016**, *60*, 177–192. [CrossRef]
86. Liu, Y.; Zhang, X. Intrusion Detection Based on IDBM. In Proceedings of the International Conference on Dependable, Autonomic and Secure Computing, Auckland, New Zealand, 8–12 August 2016; pp. 173–177.
87. Masduki, B.W.; Ramli, K.; Saputra, F.A.; Sugiarto, D. Study on Implementation of Machine Learning Methods Combination for Improving Attacks Detection Accuracy on Intrusion Detection System (IDS). In Proceedings of the International Conference on Quality in Research (QiR), Lombok, Indonesia, 10–13 August 2015; pp. 56–64.
88. Moon, D.; Pan, S.B.; Kim, I. Host-Based Intrusion Detection System for Secure Human-Centric Computing. *J. Supercomput.* **2015**, *72*, 2520–2536. [CrossRef]
89. Mukkamala, S.; Sung, A.H. A Comparative Study of Techniques for Intrusion Detection. In Proceedings of the IEEE International Conference on Tools with Artificial Intelligence (ITCAI), Sacramento, CA, USA, 3–5 November 2003; pp. 570–577.
90. Mukkamala, S.; Sung, A.H.; Abraham, A. Intrusion Detection Using Ensemble of Soft Computing Paradigms. In *Intelligent Systems Design and Applications*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 239–248.
91. Mukkamala, S.; Sung, A.H.; Ribeiro, B.M. Model Selection for Kernel Based Intrusion Detection Systems. In *Adaptive and Natural Computing Algorithms*; Springer: Vienna, Australia, 2005; pp. 458–461.
92. Mukkamala, S.; Sung, A.H.; Abraham, A.; Ramos, V. Intrusion Detection Systems Using Adaptive Regression Spines. In *Enterprise Information Systems*; Springer: Dordrecht, The Netherlands, 2004; pp. 211–218.
93. Nguyen, H.H.; Harbi, N.; Darmont, J. An Efficient Local Region and Clustering-Based Ensemble System for Intrusion Detection. In Proceedings of the 15th Symposium on International Database Engineering & Applications (IDEAS'11), Lisbon, Portugal, 21–27 September 2011; Association for Computing Machinery: New York, NY, USA, 2011; pp. 185–191.
94. Onoda, T. Probabilistic Models-Based Intrusion Detection using Sequence Characteristics in Control System Communication. *Neural Comput. Appl.* **2015**, *27*, 1119–1127. [CrossRef]
95. Ou, X.; Rajagopalan, S.R.; Sakthivelmurugan, S. An Empirical Approach to Modeling Uncertainty in Intrusion Analysis. In Proceedings of the Annual Computer Security Applications Conference, Honolulu, HI, USA, 7–11 December 2009; pp. 494–503.
96. Palenzuela, F.; Shaffer, M.; Ennis, M.; Gorski, J.; McGrew, D.; Yowler, D.; White, D.; Holbrook, L.; Yakopcic, C.; Taha, T.M. Multilayer Perceptron Algorithms for Cyberattack Detection. In Proceedings of the IEEE National Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS), Dayton, OH, USA, 26–29 July 2016; pp. 248–252.
97. Pathak, V.; Ananthanarayana, V.S. A Novel Multi-Threaded K-Means Clustering Approach for Intrusion Detection. In Proceedings of the IEEE International Conference on Computer Science and Automation Engineering, Beijing, China, 25–27 May 2012; pp. 757–760.
98. Qin, H.; Riehle, K.; Zhao, H. Using Google Analytics to Support Cybersecurity Forensics. In Proceedings of the IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 3831–3834.
99. Roshan, S.; Miche, Y.; Akusok, A.; Lendasse, A. Adaptive and Online Network Intrusion Detection System using Clustering and Extreme Learning Machines. *J. Frankl. Inst.* **2018**, *355*, 1752–1779. [CrossRef]

100. Ruiz-Agundez, I.; Penya, Y.K.; Bringas, P.G. Optimal Bayesian Network Design for Efficient Intrusion Detection. In Proceedings of the International Conference on Human System Interaction, Rzeszow, Poland, 13–15 May 2010; pp. 444–451.
101. Sabar, N.R.; Yi, X.; Song, A. A Bi-Objective Hyper-Heuristic Support Vector Machines for Big Data Cyber-Security. *IEEE Access* **2018**, *6*, 10421–10431. [CrossRef]
102. Seelammal, C.; Devi, K.V. Computational Intelligence in Intrusion Detection System for Snort Log using Hadoop. In Proceedings of the International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kumaracoil, India, 16–17 December 2016; pp. 642–647.
103. Shalaginov, A.; Franke, K. Automated Intelligent Multinomial Classification of Malware Species using Dynamic Behavioral Analysis. In Proceedings of the Annual Conference on Privacy, Security and Trust (PST), Aukland, New Zealand, 12–14 December 2016; pp. 70–77.
104. Singh, N.; Olinsky, C. Demystifying Numenta Anomaly Benchmark. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 1570–1577.
105. Singh, R.; Kumar, H.; Singla, R.K. An Intrusion Detection System using Network Traffic Profiling and Online Sequential Extreme Learning Machine. *Expert Syst. Appl.* **2015**, *42*, 8609–8624. [CrossRef]
106. Singh, S.; Silakari, S. An Ensemble Approach for Cyber Attack Detection System: A Generic Framework. In Proceedings of the ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, Honolulu, HI, USA, 1–3 July 2013; pp. 79–84.
107. Song, J.; Takakura, H.; Okabe, Y.; Nakao, K. Toward a More Practical Unsupervised Anomaly Detection System. *Inf. Sci.* **2013**, *231*, 4–14. [CrossRef]
108. Sreenath, M.; Udhayan, J. Intrusion Detection System using Bagging Ensemble Selection. In Proceedings of the IEEE International Conference on Engineering and Technology (ICETECH), Coimbatore, India, 20 March 2015; pp. 1–4.
109. Strasburg, C.; Basu, S.; Wong, J. A Cross-Domain Comparable Measurement Framework to Quantify Intrusion Detection Effectiveness. In Proceedings of the Annual Cyber and Information Security Research Conference on (CISRC), Oak Ridge, TN, USA, 5–7 April 2016; pp. 1–8.
110. Stroeh, K.; Madeira, E.R.M.; Goldenstein, S.K. An Approach to the Correlation of Security Events based on Machine Learning Techniques. *J. Internet Serv. Appl.* **2013**, *4*, 7. [CrossRef]
111. Subbulakshmi, T. A Learning-Based Hybrid Framework for Detection and Defense of DDoS Attacks. *Int. J. Internet Protoc. Technol.* **2017**, *10*, 51. [CrossRef]
112. Sukhanov, A.V.; Kovalev, S.M.; Stýskala, V. Advanced Temporal-Difference Learning for Intrusion Detection. *IFAC-PapersOnLine* **2015**, *48*, 43–48.
113. Sun, L.; Mu, W.-S.; Qi, B.; Zhou, Z.-J. A New Privacy-Preserving Proximal Support Vector Machine for Classification of Vertically Partitioned Data. *Int. J. Mach. Learn. Cybern.* **2015**, *6*, 109–118. [CrossRef]
114. Varshovi, A.; Rostamipour, M.; Sadeghiyan, B. A Fuzzy Intrusion Detection System based on Categorization of Attacks. In Proceedings of the Conference on Information and Knowledge Technology (IKT), Shahrood, Iran, 28–30 May 2014; pp. 50–55.
115. Vasudevan, A.R.; Selvakumar, S. Local Outlier Factor and Stronger One Class Classifier Based Hierarchical Model for Detection of Attacks in Network Intrusion Detection Dataset. *Front. Comput. Sci.* **2016**, *10*, 755–766. [CrossRef]
116. Vidal, J.M.; Orozco, A.L.S.; Villalba, L.J.G. Adaptive Artificial Immune Networks for Mitigating DoS Flooding Attacks. *Swarm Evol. Comput.* **2018**, *38*, 94–108. [CrossRef]
117. Vinayakumar, R.; Soman, K.P.; Poornachandran, P. Applying Convolutional Neural Network for Network Intrusion Detection. In Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI), Manipal, India, 13–16 September 2017; pp. 1222–1228.
118. Vinayakumar, R.; Soman, K.P.; Poornachandran, P. Evaluating Effectiveness of Shallow and Deep Networks to Intrusion Detection System. In Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI), Manipal, India, 13–16 September 2017; pp. 1282–1289.
119. Viswanathan, A.; Tan, K.; Neuman, C. Deconstructing the Assessment of Anomaly-based Intrusion Detectors. In *Research in Attacks, Intrusions, and Defenses Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8145, pp. 286–306.
120. Wang, Z. Deep Learning-Based Intrusion Detection with Adversaries. *IEEE Access* **2018**, *6*, 38367–38384. [CrossRef]

121. Xu, L.; Zhan, Z.; Xu, S.; Ye, K. An Evasion and Counter-Evasion Study in Malicious Websites Detection. In Proceedings of the IEEE Conference on Communications and Network Security, San Francisco, CA, USA, 29–31 October 2014; pp. 265–273.
122. Yao, F.; Yerima, S.Y.; Kang, B.; Sezer, S. Continuous Implicit Authentication for Mobile Devices based on Adaptive Neuro-Fuzzy Inference System. In Proceedings of the International Conference on Cyber Security and Protection of Digital Services (CyberSecurity), London, UK, 16–20 June 2017; pp. 1–7.
123. Yi, X.; Rao, F.-Y.; Bertino, E.; Bouguettaya, A. Privacy-Preserving Association Rule Mining in Cloud Computing. In Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security (ASIA CCS '15), Singapore, 14–17 April 2015; Association for Computing Machinery: New York, NY, USA, 2015; pp. 439–450.
124. Yin, X.; Zhang, Y.; Chen, X. A Binary-Classification Method Based on Dictionary Learning and ADMM for Network Intrusion Detection. In Proceedings of the International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), Nanjing, China, 12–14 October 2017; pp. 326–333.
125. Yu, S.; Dasgupta, D. An Effective Network-Based Intrusion Detection using Conserved Self Pattern Recognition Algorithm Augmented with Near-Deterministic Detector Generation. In Proceedings of the IEEE Symposium on Computational Intelligence in Cyber Security (CICS), Paris, France, 12–13 April 2011; pp. 17–24.
126. Zanero, S. ULISSE, A Network Intrusion Detection System. In Proceedings of the 4th annual workshop on Cyber security and information intelligence research: Developing strategies to meet the cyber security and information intelligence challenges ahead (CSIIRW'08), Oak Ridge, TN, USA, 12–14 May 2008; Association for Computing Machinery: New York, NY, USA, 2008; pp. 1–4.
127. Zhang, H.; Yao, D.; Ramakrishnan, N.; Zhang, Z. Causality Reasoning about Network Events for Detecting Stealthy Malware Activities. *Comput. Secur.* **2016**, *58*, 180–198. [CrossRef]
128. Zhang, S.; Li, B.; Li, J.; Zhang, M.; Chen, Y. A Novel Anomaly Detection Approach for Mitigating Web-Based Attacks against Clouds. In Proceedings of the IEEE International Conference on Cyber Security and Cloud Computing, New York, NY, USA, 3–6 November 2015; pp. 289–294.
129. Zhang, Y.; Wang, L.; Sun, W.; Green, R.C.; Alam, M. Artificial Immune System-Based Intrusion Detection in a Distributed Hierarchical Network Architecture of Smart Grid. In Proceedings of the IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 24–28 July 2011; pp. 1–8.
130. Zhuo, X.; Zhang, J.; Son, S.W. Network Intrusion Detection using Word Embeddings. In Proceedings of the IEEE International Conference on Big Data, Boston, MA, USA, 11–14 December 2017; pp. 4686–4695.
131. Çatak, F.Ö. Secure Multi-party Computation Based Privacy Preserving Extreme Learning Machine Algorithm Over Vertically Distributed Data. In *Neural Information Processing Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2015; pp. 337–345.
132. Lin, K.-P.; Chen, M.-S. On the Design and Analysis of the Privacy-Preserving SVM Classifier. *IEEE Trans. Knowl. Data Eng.* **2011**, *23*, 1704–1717. [CrossRef]
133. Saad, S.; Traore, I.; Ghorbani, A.; Sayed, B.; Zhao, D.; Lu, W.; Felix, J.; Hakimian, P. Detecting P2P Botnets through Network Behavior Analysis and Machine Learning. In Proceedings of the Annual International Conference on Privacy, Security and Trust (PST), Montreal, QC, Canada, 19–21 July 2011; pp. 174–180.