Cyber Security Using Machine Learning Techniques

Manisha A. Manjramkar(✉) and Kalpana C. Jondhale

MGM's College of Engineering, Nanded, India
{manjramkar_ma,jondhale_kc}@mgmcen.ac.in

Abstract. Machine learning (ML) is a subfield of Artificial Intelligence (AI) that contributes to the development of systems that can learn from previous data, spot patterns, and make logical judgments with little human interaction. Cybersecurity methodologies provide modern security solutions for detecting and responding to threats. As a result of thieves' ability to circumvent traditional security measures, the previously utilized security solutions are no longer enough. Protecting digital systems from hostile assaults, including those on computers, servers, mobile devices, networks, and associated data, is the practice of cyber security. Accounting for cyber security where machine learning is used and using ML to enable cyber security are the two main components of combining cyber security and ML. We may benefit from this union in a number of ways, including by giving ML models better security, enhancing the effectiveness of cyber security techniques, and enabling the efficient detection of zero-day threats with minimal human involvement. We combine cyber security and ML to address two distinct themes in this survey article. By providing ML strategies for cyber security, the purpose of this paper is to give a wide overview of ML methods employed in cyberspace security.

Keywords: Cyber security · Artificial Intelligence · Intrusion detection · Malware · spam

1 Introduction

The amount of time spent on the Internet has significantly grown because to advancements in computer system, internet and smart phone. Millions of various networked computers, networks, and related devices make up the global Internet. As a result, online criminals and adversaries now have the Internet as a target. Information confidentiality, availability, and integrity must all be guaranteed via a solid, secure computer system. When an unauthorized individual, software, or unlawful breach accesses a system or network with the aim to cause harm or interfere with regular operations, the computer system's authenticity and privacy are seriously compromised [1]. Cybersecurity refers to a set of safeguarding practices that may be used to secure the digital environment and user activities against unwanted access and assaults.

S. Tamane et al. (Eds.): ICAMIDA 2022, ACSR 105, pp. 680–701, 2023.
https://doi.org/10.2991/978-94-6463-136-4_59

A cyber-defense system's primary goal is to ensure data security, integrity, and accessibility [2]. Internal loopholes in computer system and network setup and implementation render them at risk of cyber-attacks and threats Weaknesses in the design of information network systems involve improper design, an absence of adequate protocols, and inexperienced or unskilled staff. These shortcomings increase the risk of threats and assaults on your network either from the inside or outside. Quite a few people from various fields are addicted to cyber networks. An agent that alters the behavior and operations of a computer or network negatively and unintentionally using a specific penetration technique is referred to as a threat [3]. Cybersecurity strives to defend against online threats to the confidentiality of data, networks, and algorithms. Between cybercriminals and defense, there has been a competition since the first computer virus emerged in 1970 [4]. It is getting harder and harder to defend against these cybersecurity threats and stay up with their pace. In order to overcome these security difficulties, researchers are currently concentrating on the urgent need to discover new automated security approaches.

One of the most efficient and beneficial methods to achieve this aim is to use independent ML algorithms to identify novel and undiscovered crimes [5]. With the use of machine learning techniques, we can identify spam, detect fraud, malware, phishing, dark or deep web sites, and uncover breaches. Human deficiencies in these particular cybercrime detection approaches can be alleviated by ML techniques. Additionally, detecting and responding to new generation cyberattacks requires a proactive approach.

One of the potential methods to rapidly resist such attacks is machine learning (ML), which has the ability to learn from experience and react to future threats in a timely manner. Today's common cybersecurity technologies include firewalls, antivirus software, intrusion prevention systems (IPS), SEIM solutions, and unified threat management (UTM). Traditional solutions rely on static management of devices in accordance with predetermined network security rules and lack automation (using AI approaches). In terms of performance, error rate, and reaction to cyberattack, AI-based systems perform better than conventional threat detection approaches. They also have lower error rates while detecting and reacting to assaults than conventional systems. ML models play a key role in improving performance and providing robust and intelligent techniques to detect attacks early and mitigate the impact and damage they cause. It combines ML techniques in order to increase precision of accurate and early cyberattack categorization. However, the majority of investigations have used inadequate datasets. Neither of the research emphasized a complete and accurate perspective of cyberattacks and threats against both computer networks and mobile devices.

2 Fundamentals of Cyber Security

As a research effort, cybersecurity's history began. In the 1970s, Robert Thomas, a researcher at Cambridge, Massachusetts-based BBN Technologies, invented the first computer "worm." Its name was Creeper. With the phrase "I'M THE

CREEPER: CATCH ME IF YOU CAN." The Creeper infected computers by bouncing from system to system. The Reaper was the first antivirus program developed by Ray Tomlinson, who invented email. It was a replicating program that would hunt for and eliminate Creeper. Robert Morris had an idea toward the end of 1988 to gauge the size of the internet. He created software to accomplish this that entered UNIX terminals, traversed networks, and cloned itself. Because the Morris worm was so aggressive, computers were rendered completely inoperable. He was later the first to be found guilty under the Computer Fraud and Abuse Act. Techniques and procedures created to protect electronic data are referred to as cybersecurity. Data is what criminals ultimately desire, after all. Computers, servers, and networks are only tools for accessing data. Effective cybersecurity reduces the possibility of cyberattacks and protects individuals and organizations against unauthorized use of technology and systems.

2.1 Attacks and Threats

The potential risks and dangers of all security breaches mentioned are called threats, and attempts to commit breaches are called attacks [8]. There are various ways to describe cybersecurity, including defining it in terms of the most dangerous assaults, such as phishing and malware [7]. Phishing, often referred to as brand cloning, is the practice of gaining access to personally identifiable information in order to manipulate or abuse it by pretending to be an authorized user.

An idea of phishing is pretending to be a scammer by using a real girlfriend's website to get personal data [9,10]. Worms, Trojan horses, and viruses are the three basic types into which malware falls. A virus is a piece of software that degrades a computer's performance without the user's awareness. Viruses can harm the operating system and files on your computer. In 1981, Elk Cloner was the inaugural computer to propagate through floppy disc drives [11]. A worm is a computer software that replicates itself continuously while using up system or network resources. Unlike viruses and worms, Trojan horses masquerade as genuine software and are launched by certain procedures or actions instead of multiplying itself. Unwanted spam emails are another danger to cybersecurity. Not only do these emails take a long time to fill up your inbox, but they are also the source of the Java applets that run while you read your emails. Spam can appear as calls, texts, and video communications on mobile devices and networks [12–15]. On Twitter and YouTube, spammers frequently target text messages and videos. Firewalls, antivirus software, and intrusion detection systems are all components of network security systems. Unauthorized intrusions and malicious illegal access are detected and identified with the use of intrusion detection systems (IDS).

3 Fundamentals of ML

To enhance cybersecurity, identify phishing websites, and identify numerous automated new assaults early on, ML approaches are deployed [22]. In terms of approach, ML may be split into three primary groups: semi-supervised, unsupervised,

and supervised. The machine already knows the targeted labeling or classes of data in supervised machine learning, and the machine uses these labeling and classes to train the computer. Unsupervised ML does not provide the intended value. The primary goal of unsupervised learning is to find data associations. It acts by looking for patterns in data, like in Clustering. Semi-supervised machine learning (ML) refers to a method where some of the data is labelled or when human expertise is needed during data collecting. During the lettering process, human specialists can assist with problem- solving and enhancing model accuracy [6].

A different branch of machine learning is reinforcement learning (RL). RL is also known as "learning through criticism." This is due to the fact that each inaccurate forecast acts as an input to the algorithm. However, the method was not told how to rectify it. Instead, the method must find and test several interpretations until it determines the successful outcome. A part of machine learning is deep learning (DL). While ML and deep learning use similar methods and tackle similar problems, they are not equivalent in terms of power. The way the human brain processes information motivates DL algorithms to do the same. Deep Learning (DL) comprises two primary areas of research: conventional neural networks and deep belief networks [25–28]. DL models are used in several researches in the literature to enhance cybersecurity [29,31].

3.1 Popular Machine Learning Technologies

This paper explains typical machine learning approaches. Table 1 shows frequently used ML techniques, running time, benefits and drawbacks and announced year.

Support Vector Machine (SVM). When it comes to IDS, SVMs are thought to be the most well-liked and effective ML approach. Based on the labeling of the margins on either side of the hyperplane, the SVM classifies and divides the two data classes. Figure 1 is a visual representation of SVM. The gap between margins and hyperplanes can be increased to improve classification results. Support vector points are data points located at the hyperplane's boundary. The supervised learning algorithm known as the SVM belongs to the category of classification techniques. This method of binary categorization forecasts the ideal hyperplane in n-dimensional space using a training data set. Data in two-dimensional planes and multidimensional hyperplanes are classified using the SVM algorithm. A multidimensional hyperplane classifies multidimensional data using a "kernel." Maximum discriminative power between categorized data points is always preferred. In other terms, the hyperplanes' maximal margins between data points, or their maximal spacing, should be used. A hyperplane is a boundary that divides a plane. SVMs are used to classify multidimensional data, so the hyperplane is a straight line with two inputs and a 2D plane with three or more inputs.

The SVM algorithm is primarily used for classification, although it can also be used for regression analysis. In order to anticipate the outcome, a classification algorithm examines the training data. The result is a class, such as Day or Night,

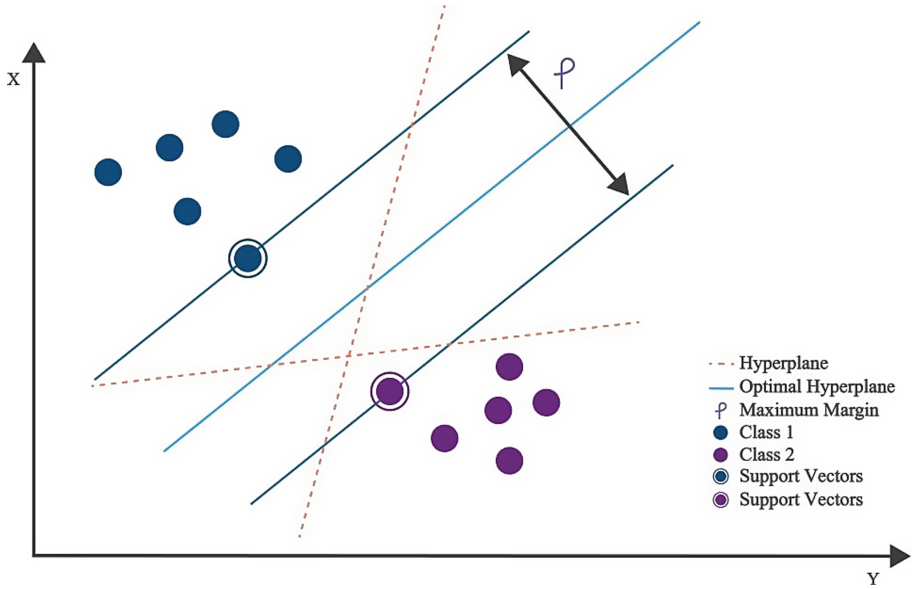

Fig. 1. Support Vector Machine

Yes or No, Long or Short. When a buyer in a shop buys bread and butter at the same time, this is an example of a classification algorithm. The response for the target class would be either Yes or No.

The association between the independent variables is discovered via a regression method, which also forecasts the result. SVMs are divided into two distinct categories. It could be linear or nonlinear, depending on the kernel function. It might be one or many classifications, depending on the type of recognition [34,35]. SVMs require a lot of memory to process and take a long time to train. To achieve improved outcomes for learning dynamic usage patterns, SVM should be trained with various time intervals. Kernel functions and parameters also affect classifier performance.

Decision Tree (DT). DT are frequently used in ML to categorize objects based on previously learned features. The method may be used to correct regression errors or to forecast consistent results from unknown data. The key advantages of employing a decision tree in ML are ease of interpretation and understanding of the decision-making process. Because decision trees in ML can result in excessively complex branches, pruning of the tree structure is frequently necessary. Using decision trees to model decisions and outcomes is one way to map decisions in a branching structure.

As shown in Fig. 2, DT consists of root or intermediate nodes, paths and leaf nodes. The root or intermediate nodes of the tree represent objects or attributes. Every branching path in the tree reveals a possible parent node value. Leaf nodes

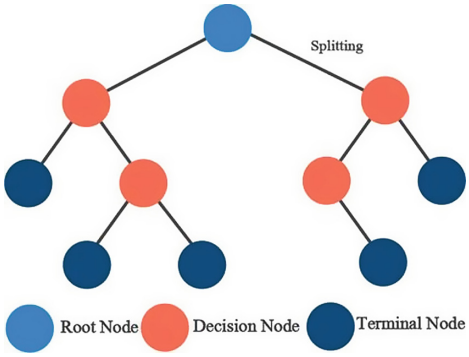

Fig. 2. Decision Tree

relate to expected categories or classified properties. If-then rules are another way to express the resultant tree. We then use diversity and collect more information parameters to select the best routing information as we build the tree. CART [36], C4.5 [37], and ID3 [38] are important methodologies in DT. C4.5 addresses the ID3 limitations by using a tree pruning strategy to cope with the overfitting problem.

K-Nearest Neighbor (KNN). KNN is an algorithm for unsupervised classification. It is based on a distance function, which calculates the difference between two data instances. It requires minimum time to train other classifiers. However, it's computing time is consumed during the identification step. Figure 6 demonstrates how KNN works. This classifier is based on the concept that comparable data points in the region will be nearer to each other than different data points. Based on anomaly ratings, there are two primary groups of KNN. The two methods for computing anomaly ratings are as follows:

1. Calculation based on the variance between the data point and the kth neighbor.
2. Calculation depending on each data instance's frequency [39].

The kth data point's value has an impact on the classifier's overall performance [40]. The choice of distance function to determine the separation/difference between data points and noisy data both affect how well the classifier performs (Fig. 3).

KNN is costly to compute and needs a lot of storage to be manipulated. The distance among data points x and y is commonly determined using the Euclidean distance formula, $d(x, y)$.

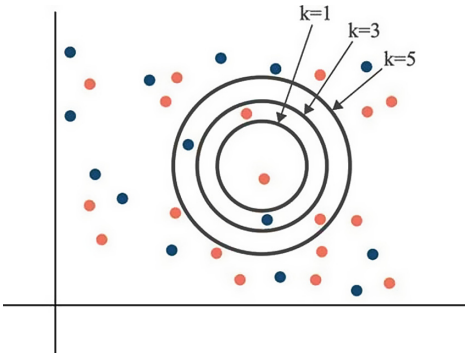

Fig. 3. k-Nearest neighbour

Random Forest (RF). RF is a form of ensemble learning that uses numerous classification models to generate a scientific consensus about an issue in order to build up a typical results. RF is viewed as a more advanced version of CART. An RF typically contains a number of forecasting outcomes derived from several decision trees. In literature, random forest is used for issues like intruder identification and examination spam quantity [41,42].

It enhances nonlinear problem performance and uses less processing power during in the model training stage. The decision tree algorithm that should be taken into account during the prediction step must be chosen, though, because Random Forest can predict numerous decision trees.

Naive Bayes (NB). NB is a group of classifiers that breaks down the conditional probabilities of the topic under study and is based on Bayesian' theorem. This independence criterion does not, however, applicable to certain sorts of assaults in cybersecurity. An improved version, called Hidden NB, may resolve these issues with an accuracy rate of 99.6% [43]. With discrete characteristics, NB classifiers perform well. This classifier is thought to be easier to identify and faster. Polynomial, Bernoulli, and Gaussian are three of NB's effective approaches. Distinct inputs are managed by multinomial NB. A feature vector of these values represents the recurrence of this issue. Binary feature vector classification is done using Bernoulli NB. A classifier for continuous data values is Continuous Gaussian NB. The Gaussian distribution is used to describe the distribution of these numbers [47].

Artificial Neural Network (ANN). Both forward and reverse propagating cycle's method is used to train ANNs. Each buried layer node receives data via the feed forward method. It computes the initiation factor for every concealed and output layer node. Performance of classifiers is influenced by activation functions. By comparing the network output to the target value, error is computed. In order to change the values between hidden and output nodes, backpropagation

algorithm transmits the change corresponding to the input layer and applies the Guardian Descent algorithm. Until the required level is attained, this process is repeated [45].

ANNs are simple to use, noise-resistant, and nonlinear models, but they are slow to train. Taveras's [46] goal was to examine how crucial end-user password entering habits are to account security. To reduce the chance of account hacking, they advise changing the password entering habits. They asked individuals to write down any passwords as part of their investigation. To predict future outcomes in this analysis, ML methods, notably artificial neural, were employed. Overall, the study indicated that while neural networks may be utilized for prediction quite well, they still have significant drawbacks. Due in part to the fact that the majority of the participants had backgrounds in information technology, the user's activities did not always proceed logically. Robust data collection improves model precision and assists in the search for difficulties caused by end-user credential entry practices.

Recurrent Neural Network (RNN). RNNs are a type of neural network. Invisible states exist in RNN [44]. As illustrated in Fig. 4 (a), every state receives the result of the preceding stage as its input. Inside this method, data travels across RNN phases. The primary purposes of RNN are to analyze time series data and to construct streaming data processing. RNNs have memory, which allows them to recall information from prior knowledge and utilize it as input for succeeding states [47].

Convolutional Neural Network (CNN). CNN is a feed-forward ANN extension that is multi-layer neural network. It is composed of three types of layers, as illustrated in Fig. 4 (b): a convolutional layer or layers, one or more fully interconnected layers, and pooling layers. CNN architectures such as GoogLeNet [48] and ResNet [49] are widely employed. Extraction of complex high-resolution features accompanied by the transformation into finely ground complicated features.

CNNs are used in a wide range of areas, including drug research, defect detection and picture categorization. Using the KDD99 dataset, Riaz et al. suggested an improved version of CNN with 99.23% accurate intruder detection [50]. CNNs are frequently employed to categorize harmful traffic. Deep neural networks (DNNs) are employed in airline passenger profiling, identifying normal travelers and prospective attackers.

Deep Belief Network (DBN). DBN is a deep neural network subset that uses an unsupervised optimization technique. DBNs were created to understand the human mind's ability to absorb complex data and uncover complex patterns. DBN is a stack of Restricted Boltzmann Machine (RBM) with crucial generative character. Unlike RBM, however, DBN does not interact across nodes in the same network layer. Every node inside the deep learning model is linked to all nodes in the previous and subsequent levels. A DBN is fed data in the form of probabilities.

A DBN needs each network layer to learn a whole input and provide an output [51]. Each layer continues to provide optimum options at each step, which is continued until the training stage is finished to the required level, as shown in Fig. 4(c).

Reinforcement Learning (RL). Another subfield of ML is RL. RL is also described as gaining knowledge with a reviewer since the algorithms receive feedback for any inaccurate predictions. However, the algorithm has not been informed how to fix it. Instead, the algorithm must work out and test various alternatives until it discovers the proper solution [23]. This process operates on a reward and punishment system. Deep learning methods are applied to handle a variety of complex problems. AlphaGo [24] is an example of this method. Deep RL is implemented in cybersecurity for identifying host assaults, reducing Assaults [53], identifying phishing emails [54], and cyber- physical systems [55], among other things. RL is supposed to be the closest method to replicating how human thinking is regarded to occur by leveraging the unknown and unfamiliar environment.

Figure 7 demonstrates, RL functioning is made up of five parts: agent, surroundings, award, status, and benefit. Agents create their own learning experiences by interacting with their immediate surroundings. This action has resulted

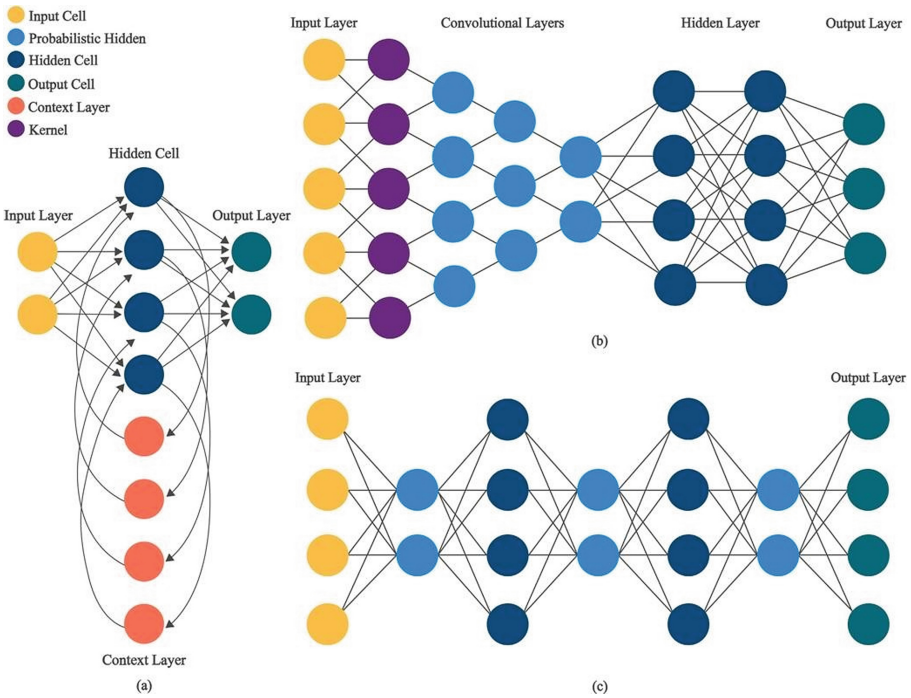

Fig. 4. (a) RNN (b) CNN (c) DBN.

in two modifications. First, the ambient condition is altered to reflect the new status. Second, based on activities, the surroundings imposed a fine or incentives. The reward function, given a state, informs the agent how successfully or badly the attempt accomplished. The individual learns from incentives and filters out unwanted behavior.

4 Cyber Security Using ML

Cybersecurity guarantees safeguarding against cyber dangers. Cybersecurity encompasses various dimensions, such as harmful URL detection and categorization, fraudulent transactions, spam classifications, intrusion detection systems (IDS), malicious nodes creation, probing, cyber-extortion, and malware. Furthermore, as computer networks have developed, just have smart devices and networks, rendering them a target for cybercriminals. Cybersecurity connects with other cyberspace elements including such internet safety, network security and ICT security. Three important cybersecurity concerns (IDS, spam, and malware recognition and identification) have been addressed, with ML technology playing a significant role.

In computer networks, intrusion detection systems would be furthermore subdivided into signature-based, exploit-based, anomaly-based, and hybrid-based methodologies. Intrusion subgroups are further classified into those that affect hosts or computer networks. Spam detection is examined in greater detail with respect to media such as photos, emails, SMS, videos, and Twitter. Malware is also investigated via static and dynamic analysis. In the previous research, ML approaches have been used to counteract several sorts of cyber-attacks. One of the tools that can swiftly fight cyber threats is machine learning. ML approaches are applied to challenges where learning methods may learn from experience and respond fast to new threats.

The subsections that follow clarify on each cyber danger to the computer system and cellular technologies, as well as how cutting-edge ML approaches are being used to combat these cyber-attacks.

4.1 Spam Detection by Using ML

Electronic mail is a technique of transferring data among people utilizing digital devices via the Internet. It is extensively employed as a tool and it has rapidly grown in popularity. Spam messages are unnecessary, unwanted emails that are regularly utilized for advertising and irritate or bother consumers. Spam email costs traffic and disc storage; it significantly lowers the duration and time spent online, as well as the functionality of networks and systems [51]. Nowadays, more than 85% of all emails or conversations received are spam [16]. Spammers find their email as well as online search engines to be prime targets. Email spam is not the only item that's been impacted. Spam is on the rise across a variety of mediums, including smart phones, blogging, newsletters, online chatting, telephone conversations, and streaming sites. Social media networks such as Facebook,

Twitter, and YouTube allow scammers to easily upload and spread material, which scammers take advantage of them. Computer researchers are looking for a rapid and essential answer for this. Spam filtration is the process of identifying email as ham or spam and screening out undesirable email [47]. In the literature, several spam filtering strategies have been suggested. However, it is inefficient since spammers are intelligent enough to modify spam terms. Anti-spam, often known as anti-spam technology, is a collection of steps implemented to combat various spam attacks while minimizing the impact on the effectiveness of the intended medium.

ML approaches are being developed to increase efficiency and combat spammer assaults. Many machine learning (ML) algorithms for spamming categorization, filtration, and detection have really been reported in the literature. Machine learning approaches are employed in a variety of spam detection domains, including Twitter, picture emails, and blogging. Every area has its own classification technique. Most research, meanwhile, suggest that the SVM approach is more successful than alternative classifications. Numerous writers used a feature selection approach accompanied by a classifier to considerably increase the classifier's accurateness. Furthermore, using several classifiers to increase accuracy rate might be a future study topic. Decision Trees, J48, NB, SVM, and Random Forest are popular ML approaches. Signature-based approaches are conventional spam detection methods that employ signatures to recognize potential harm. Despite this, the detection rate for fresh spam assaults is poor [62]. Many email systems have filtering features, but users may acquire additional security and control by purchasing filtering software. Methods such as content based, machine learning [63], and quarantining [64] are also employed to accomplish the same outcome. It offers a number of spam filtering tools and strategies. Datasets from Spam base, Enron, PU Datasets, and Ling-Spam are extensively utilized spam categorization and filtration [57–61]. Email is often regarded as a popular entry route for viruses. Hitting on a hyperlink in an email by mistake might place your machine and connection at danger. Because emails and papers might include a large number of lines, the feature space is limited.

Feature selection refers to the process of selecting the best subgroup from the most key features. Feature selection improves the accuracy and usefulness of the training and classification procedures significantly. For spam detection, J48, Bayes Net, and SVM were evaluated, with SVM surpassing the others. In addition, J48 excelled SVM when it came to spam mail categorization [65]. The increase in People on twitter has led to an increase in spam tweets. Spam Tweets is uninvited and uninvited Tweets that include harmful code and can lead to further privacy issues including phishing, forgery, selling drugs, and malware transmission. Various machine learning approaches are used to analyze streaming spam tweets [98]. According to the survey, NB fared better with 97.3% accuracy. The authors used decision tree, random forest, and NB approaches, with decision tree algorithm achieving the greatest accuracy [61].

Spam detection technologies are divided into two group's content assessment and image-based assessment. Image-based spam, which spammer's target,

is undetectable by textual analyses. To identify image-based spam, several pattern recognition and computer vision algorithms are applied. SVM is a popular machine learning approach for identifying spam in blogs [66,67], and videos [68]. For spam detection in blogging and videos, decision trees, Firefly, and Bays classifiers are utilized.

There are a variety of anti-spam technologies available to safeguard consumers against trash and unsolicited emails. Solar Winds MSP Mail Assure [69], Spam Titan [70], SPAMfighter [71], and ZEROSPAM [72] are among the anti-spam methodologies.

Smartphone and services are getting increasingly popular these days. Spammers target smartphones such as Messaging, email applications, pictures, data, cellular cloud, and phone calls. SMS is viewed as a basic and low-cost method of

Table 1. Frequently used ML techniques

Model	Year	Time Complexity	Description	Limitations
SVM	1995	$O(n^3)$	• Can be used for classification and regression. • Less overfitting	• Unable to handle large or noisier datasets efficiently. • High computational cost.
Naive Bayes	1960	$O(mn)^2$	• A probabilistic classifier that takes less computational time. • Assumes that a feature is entirely independent of all other present features.	• Assigns 0 probability if some category in the test data set is not present in the training data set. • Stores entire training examples • Need massive data to obtain good results.
Random forest	1995	$O(M\ m \log n)^2$	• Composed of many DTs. • Every DT yields a prediction. • The prediction having a maximum number of votes will be the final prediction of the model.	• Computational cost is higher. • Slow prediction generator
ANN	2000	$O(amnk)^2$	• Adaptive and composed of interconnected Artificial Neurons. • Next Layer input depends on Previous Layer Output.	• High cost and time-consuming. • Black-box model hence shows no relation between input and output variable.
Decision Tree	1979	$O(mn^2)^5$	• Works on an if-then rule to find the best immediate node. • Continue the process until the predicted class is obtained	• Difficult to change the data without affecting the overall structure. Complex, expensive and time consuming
K-Mean	1960	O(kmni)^6	• Starts from random centroids refine centroids in iterations till the final cluster analysis.	• High dependency on initial centroids. inefficient clustering for varying cluster sizes
DBN	2006	$O\left(m\sum_{i=1}^{l}(l_i l_i)\right)^7$	• Higher performance and efficiency is achieved because of the addition of the layers. • Better ability to handle noisy data. • Convenient identification of complex relationships between nodes. • Hidden layers are efficiently used.	• Higher hardware resources consumption. • Higher time consumption because of the addition of the layers. • Unable to provide an explanation for the decisions.
RNN	1982	-	• Efficiently models sequential data. • Quickly memorize the sequential events. • Different various, i.e. LSTM are available.	• Difficult training of the network. • It may face short memory issues while • Modelling long sequences of data. • Vanishing Gradient and gradient exploding problems.
CNN	1988	$O(\sum_{l=1}^{d} n_{l-1} \cdot s_l^2 \cdot n_l \cdot m_l^2)$	• Less number of neurons are needed in contrast with traditional NN. • Different variants, e.g. VGG, AlexNet, are available.	• It requires more number of convolutional layers(CL). • A larger tagged dataset is necessary for working.

phishing attempts. Smartphones contain the data such as debit and credit card details and login details such as user credentials.

Spam SMS is commonly associated with free services, ads, promotions, bundles, and bonuses [73]. On mobile devices, ML approaches play a crucial role in spam detection and identification. SMS, phone calls, email apps, data on mobile devices, photos, and videos are all examples of spam. To detect spam, researchers used SVM, Nave Bayes, KNN, RNN, and k-means ML approaches. The Bayesian learning strategy was used.

Image sharing via various communication platforms such as Instagram, WhatsApp, Facebook, and others has grown dramatically. Many research have been conducted on the filtering and classification of spam images [75]. Malicious calls, including phone fraud and spam, have become a major issue throughout the world in recent years. The author applied SVM [76], random forest, and logistic regression to detect spam calls and minimize harmful calls by 90%.

4.2 Intrusion Detection by Using ML

There are three major types of cyber analysis for intrusion detection systems. These detections are exploit-based, anomaly-based, and hybrid-based. The purpose of exploit- based detection is to detect known attacks. Anomaly detection monitors normal network and system activity and recognizes irregular network and system behavior. Finally, to improve detection results, a hybrid-based detection technique combines exploit-based and anomaly-based strategies [6]. Attackers can successfully exploit the ubiquitous vulnerabilities of these traditional defenses. As a result, safeguarding users from new and developing threats has become difficult. There is a huge amount of data in cyber infrastructure.

It's critical to understand the patterns and behavior of invaders and assaults. As a result, ML algorithms play a major part in identifying and forecasting future intrusions and assaults in real time. ML techniques are often used to detect intruders and commonly used approaches are ANNs, fuzzy associations, SVMs, decision trees, and statistical models.

To increase the quality and recognition rate of intrusion, case-based methodology and other unsupervised learning approaches are used. Many classifiers have outperformed other classifiers in diverse areas and activities in ID. However, quick and timely identification of new and zero-day assaults remains a difficult field of research. ML algorithms were used for violence detection, anomaly detection, and hybrid detection. Cybersecurity risks in cyberspace may be divided into two types: network- based attacks and host-based threats. A cyber defense system provides safeguards at both levels. Network flow regulation is handled by network-based defense systems. In contrast, host-based defense systems communicate with firewalls and other relevant mechanisms installed on hosts to prohibit data from accessing devices [77].

Various ML techniques have been applied to detect DoS attacks, including Decision Trees with 97.24% accuracy [78], Neural Networks with 97% accuracy [79], Nave Bayes with 96.65% accuracy [78], and SVM with 91.6% accuracy [80].

SVMs and decision trees have been used to identify malicious attacks on networks [56]. RF and ANN were used to the network for hybrid intrusion detection. Various intrusion detection tools are commercially available. Intrusion detection tools are intended to manage host or network intrusions. A The Network Intrusion Detection System (NIDS) detects network intrusions. The Host Intrusion Detection System (HIDS) identifies signature-based or anomaly-based host assaults. A variety of freeware ID tools are available. Some, meanwhile, are too costly. McAfee NSP [83], Hillstone NIPS [85], and Palo Alto [86] are examples of popular commercial ID solutions. Snort, Suricata, Samhain, Security Onion, and Sagan are all free tools [78].

The tools used are determined by the operating system, detecting type (HIDS, NIDS), and detection mechanism (signature-based, anomaly-based). Another technique for preventing and mitigating cyberattacks is Trusted Automated Exchange of Intelligence Information (TAXII). TAXII uses Structured Threat Information Expression (STIX), a language designed to express cyber threat data, to indicate how facilities and communication exchanges may be employed to share threat intelligence [99].

4.3 Malware Detection by Using ML

Malicious programs also known as 'Malware' is software that is secretly installed into a device or network with the goal of compromising user's activities. Malware affects the integrity, confidentiality, and availability of data saved on hardware or software by the perpetrator. The term "malware" is derived from the words "malicious" and "ware." Viruses, worms, Trojan horses, spyware, and adware are some examples. Malicious software infects a system and spreads to other devices and networks. According to McAfee's 2019 statistics report, the total number of known malware samples has surpassed 800 million. Malware has grown considerably in recent years, causing economic loss. Malware not only targets individuals, but it also disrupts companies and the defense via skilled hackers and custom malware. Malware is recognized as the most serious potential threat to enterprises. Previously, signature-based procedures were carried out to detect malware.

ML approaches are effective not just in detecting zero-day threats, but also at detecting new or complex malware assaults. With 29% utilization, SVM is the widely researched ML classification method for malware detection, followed by Decision Trees with 17% usage [100]. Furthermore, merging DBN with other semi-supervised learning approaches boosted accuracy rate.

Malware may be divided into two generations. Malware of the initial version has a similar structure. In the 2nd version, it alters its structure and grows into a new variation while maintaining the same characteristics [88]. Based on the structural modification, the second version of malware is further categorized as encrypted, Oligormorphic, Polymorphic, and Metamorphic. Malware variations are unexpected and uncertain [101]. AI-based cyber-attacks have been detected and classified using Principal Component Analysis (PCA) and ANN.

There are several malware detecting technologies available on the market. However, selecting the proper tools is critical. Some tools are free, while others need an annual subscription cost. Avast Internet Security [88] is a popular anti-malware product, accounting for 15.21% of the overall market. Other regularly used products include Malware-Bytes, Norton Power Eraser, AVG, and Bitdefender Antivirus [90–93].

Threats to mobile devices have expanded along with the usage of mobile banking, transactions, and e-commerce. As a result, mobile devices are becoming more prone to cyber-attacks than personal computers [94].

SVM, KNN, random forests, and decision trees are some of the approaches often used to identify malware on mobile devices and networks. Classifier correctness can also be improved by feature selection preceded by classification approaches. Antimalware approaches for mobile devices are classified into three types: static, dynamic, and hybrid. Static detection is a detection approach that examines a program for harmful patterns without executing it. In contrast, dynamic detection is performed by running the real program and observing the dynamic behavior. The hybrid malware approach detects malware by combining static and dynamic analysis [95], [96].

On the model described at opcode-sequence-frequency, decision trees, KNN, and SVM were utilized to obtain 90% accuracy [105]. For malware identification, RF, SVM, and Naive Bay were also used, with Random Forest beating the others in terms of TPR and FPR [106].

5 Conclusion

As a result of their ineffectiveness in identifying previously unidentified and polymorphic attacks, the traditional security solutions that were previously in use are no longer adequate. By providing a thorough overview of the intersections between the two fields, we've narrowed the gap between different ML approaches [2] and threats to computer networks and mobile communication in this study. The literature overview on ML algorithms for malware detection, spam detection, and intrusion detection on computer networks and mobile devices over recent years is presented in this survey.

In order to combat these cybercrimes, we have created a graphic overview of the ML method now in use. With the goal of improving security measures to recognize and respond to intrusions, cyber security has grown into a global concern. ML techniques are used in numerous applications of cyber security systems in significant ways. For all attacks based on a single model, one recommendation is not conceivable. The basics of cyber security, including how to categorize intrusions on computer networks and mobile devices, have been covered. If a newcomer reads our descriptions of the ML foundations, they will have a more favorable knowledge of this field, subtypes, and important techniques due to the importance of ML. We have provided an overview of various well-known ML tools. Instead than focusing on the model's speed and accuracy, trustworthy ML uses safe ML algorithms to provide some high-level correctness.

References

1. V. Ambalavanan, "Cyber threats detection and mitigation using machine learning," in Handbook of Research on Machine and Deep Learning Applications for Cyber Security. Hershey, PA, USA: IGI Global, 2020, pp. 132-149.
2. T. Thomas, A. P. Vijayaraghavan, and S. Emmanuel, "Machine learning and cybersecurity," in Machine Learning Approaches in Cyber Security Analytics. Singapore: Springer, 2020, pp. 37-47. [Online]. Available: https://link.springer.com/chapter/10.1007/978-981-15-1706-8_3
3. F. Farahmand, S. B. Navathe, P. H. Enslow, and G. P. Sharp, "Managing vulnerabilities of information systems to security incidents," in Proc. 5th Int. Conf. Electron. Commerce (ICEC), 2003, pp. 348-354.
4. P. Szor, The Art of Computer Virus Research and Defense: ART COMP VIRUS RES DEFENSE_p1. London, U.K.: Pearson, 2005. I. Firdausi, C. Lim, A. Erwin, and A. S. Nugroho, "Analysis of machine learning techniques used in behavior-based malware detection," in Proc.2nd Int. Conf. Adv.
5. A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," IEEE Commun. Surveys Tuts. vol. 18, no. 2, pp. 1153-1176, 2nd Quart., 2016.
6. J. M. Torres, C. I. Comesaña, and P. J. García-Nieto, "Machine learn- ing techniques applied to cybersecurity," Int. J. Mach. Learn. Cybern., vol. 10, no. 10, pp. 2823-2836, 2019. Difference Between Threat and Attack. Accessed: Jun. 3, 2020. [Online]. Available: https://www.geeksforgeeks.org/difference-between-threat-and-attack/
7. S. Purkait, "Phishing counter measures and their effectiveness-literature review," Inf. Manage. Comput. Secur., vol. 20, no. 5, pp. 382-420,Nov. 2012.
8. R. M. Mohammad, F. Thabtah, and L. McCluskey, "Tutorial and critical analysis of phishing Websites methods," Comput. Sci. Rev., vol. 17, pp. 1-24, Aug. 2015.
9. E. H. Spafford, "Computer viruses as artificial life," Artif. Life, vol. 1, no. 3, pp. 249-265, Apr. 1994. H. Drucker, D. Wu, and V. N. Vapnik, "Support vector machines for spam categorization," IEEE Trans. Neural Netw., vol. 10, no. 5, pp. 1048-1054, Sep. 1999.
10. N. Jindal and B. Liu, "Review spam detection," in Proc. 16th Int. Conf. World Wide Web, 2007, pp. 1189-1190. S. M. Abdulhamid, M. S. Abd Latiff, H. Chiroma, O. Osho, G. Abdul-Salaam, A. I. Abubakar, and T. Herawan, "A review on mobile SMS spam filtering techniques," IEEE Access, vol. 5, pp. 15650-15666, 2017.
11. D. D. Arifin and M. A. Bijaksana, "Enhancing spam detection on mobile phone short message service (SMS) performance using FP-growth and naive Bayes classifier," in Proc. IEEE Asia Pacific Conf. Wireless Mobile (APWiMob), Sep. 2016, pp. 80-84. Kharon Malware Dataset. Accessed: Aug. 8, 2020. [Online]. Available: http://kharon.gforge.inria.fr/dataset/
12. D. Michie, D. J. Spiegelhalter, and C. Taylor, "Machine learning," Neural Stat. Classification, vol. 13, 1994. S. Dua and X. Du, Data Mining and Machine Learning in Cybersecurity. New York, NY, USA: Auerbach, 2016.
13. S. Angra and S. Ahuja, "Machine learning and its applications: A review," in Proc. Int. Conf. Big Data Anal. Comput. Intell. (ICBDAC), 2017, pp. 57-60.
14. T. M. Alam, K. Shaukat, M. Mushtaq, Y. Ali, M. Khushi, S. Luo, and A. Wahab, "Corporate bankruptcy prediction: An approach towards bet- ter corporate world," Comput. J., pp. 1-16, Jun. 2020.

15. A. Kulkarni and L. L. Brown, III, "Phishing websites detection using machine learning," Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 7, pp. 8-13, 2019, https://doi.org/10.14569/IJACSA.2019.0100702.
16. M. Islam and N. K. Chowdhury, "Phishing Websites detection using machine learning based classification techniques," in Proc. 1st Int. Conf. Adv. Inf. Commun. Technol., 2016, pp. 1-4. S. Marsland, Machine Learning: An Algorithmic Perspective. Boca Raton, FL, USA: CRC Press, 2014.
17. S. R. Granter, A. H. Beck, and D. J. Papke, "AlphaGo, deep learning, and the future of the human microscopist," Arch. Pathol. Lab. Med., vol. 141, no. 5, pp. 619-621, May 2017.
18. D. Yu and L. Deng, "Deep learning and its applications to signal and information processing [Exploratory DSP]," IEEE Signal Process. Mag., vol. 28, no. 1, pp. 145-154, Jan. 2011.
19. Y. Bengio, "Learning deep architectures for AI," in Foundations Trends Machine Learning, vol. 2, no. 1. Boston, MA, USA: Now, 2009.
20. R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch," J. Mach. Learn. Res., vol. 12 pp. 2493-2537, Aug. 2011.
21. P. Le Callet, C. Viard-Gaudin, and D. Barba, "A convolutional neural network approach for objective video quality assessment," IEEE Trans. Neural Netw., vol. 17, no. 5, pp. 1316-1327, Sep. 2006.
22. L. F. Maimo, A. L. P. Gomez, F. J. G. Clemente, M. G. Perez, and G. M. Perez, "A self-adaptive deep learning-based system for anomaly detection in 5G networks," IEEE Access, vol. 6, pp. 7700-7712, 2018.
23. A. Abeshu and N. Chilamkurti, "Deep learning: The frontier for dis- tributed attack detection in fog-to-things computing," IEEE Commun. Mag., vol. 56, no. 2, pp. 169- 175, Feb. 2018.
24. T. M. Kebede, O. Djaneye-Boundjou, B. N. Narayanan, A. Ralescu, and D. Kapp, "Classification of malware programs using autoencoders based deep learning architecture and its application to the microsoft malware classification challenge (BIG 2015) dataset," in Proc. IEEE Nat. Aerosp. Electron. Conf. (NAECON), Jun. 2017, pp. 70-75.
25. S. Purushotham, C. Meng, Z. Che, and Y. Liu, "Benchmarking deep learning models on large healthcare datasets," J. Biomed. Informat., vol. 83, pp. 112-134, Jul. 2018.
26. S. Feng, H. Zhou, and H. Dong, "Using deep neural network with small dataset to predict material defects," Mater. Des. vol. 162, pp. 300-310, Jan. 2019.
27. W.-H. Chen, S.-H. Hsu, and H.-P. Shen, "Application of SVM and ANN for intrusion detection," Comput. Oper. Res., vol. 32, no. 10, pp. 2617-2634, Oct. 2005. 209
28. B. Schölkopf, R. C. Williamson, A. J. Smola, J. Shawe-Taylor, and J. C. Platt, "Support vector method for novelty detection," in Proc. Adv. Neural Inf. Process. Syst., 2000, pp. 582-588.
29. A. L. Prodromidis and S. J. Stolfo, "Cost complexity-based pruning of ensemble classifiers," Knowl. Inf. Syst., vol. 3, no. 4, pp. 449-469, Nov. 2001. 211
30. J. R. Quinlan, C4. 5: Programs for Machine Learning. Amsterdam, The Netherlands: Elsevier, 2014.
31. V. H. Garcia, R. Monroy, and M. Quintana, "Web attack detection using ID3," in Proc. IFIP World Comput. Congr. (TC), vol. 12. Boston, MA, USA: Springer, 2006, pp. 323- 332.

32. V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM Comput. Surv., vol. 41, no. 3, p. 15, 2009.
33. A. A. Aburomman and M. B. Ibne Reaz, "A novel SVM-kNN-PSO ensemble method for intrusion detection system," Appl. Soft Comput., vol. 38, pp. 360-372, Jan. 2016.
34. S. He, G. M. Lee, S. Han, and A. B. Whinston, "How would information disclosure influence organizations' outbound spam volume? Evidence from a field experiment," J. Cybersecurity, vol. 2, no. 1, pp. 99-118, Dec. 2016.
35. S. T. Miller and C. Busby-Earle, "Multi-perspective machine learning a classifier ensemble method for intrusion detection," in Proc. Int. Conf. Mach. Learn. Soft Comput. (ICMLSC), 2017, pp. 7-12.
36. L. Jiang, H. Zhang, and Z. Cai, "A novel Bayes model: Hidden naive Bayes," IEEE Trans. Knowl. Data Eng., vol. 21, no. 10, pp. 1361-1371, Oct. 2009.
37. S. Sathasivam and W. A. T. W. Abdullah, "Logic learning in hopfield networks," 2008, arXiv: 0804.4075. [Online]. Available: http://arxiv.org/abs/0804.4075228
38. A. Jagota, "Novelty detection on a very large number of memories stored in a hopfield- style network," in Proc. Seattle Int. Joint Conf. Neural Netw. (IJCNN), vol. 2, 1991, p. 905.
39. P. Taveras and L. Hernandez, "Supervised machine learning techniques, cybersecurity habits and human generated password entropy for hacking prediction," in Proc. MWAIS, vol. 38, 2018, pp. 1-6. [Online]. Available: http://aisel.aisnet.org/mwais2018/38
40. J. M. Gómez Hidalgo, G. C. Bringas, E. P. Sánz, and F. C. García, "Content based SMS spam filtering," in Proc. ACM Symp. Document Eng. (DocEng), 2006, pp. 107- 114.
41. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolu- tional networks," in Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer, 2014, pp. 818-833.
42. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770-778.
43. R. U. Khan, X. Zhang, M. Alazab, and R. Kumar, "An improved convolutional neural network model for intrusion detection in net- works," in Proc. Cybersecurity Cyberforensics Conf. (CCC), May 2019,pp. 74-77.
44. G. Tzortzis and A. Likas, "Deep belief networks for spam filtering," in Proc. 19th IEEE Int. Conf. Tools Artif. Intell. (ICTAI), vol. 2, Oct. 2007, pp. 306-309.
45. X. Xu and T. Xie, "A reinforcement learning approach for host-based intrusion detection using sequences of system calls," in Proc. Int. Conf. Intell. Comput. Berlin, Germany: Springer, 2005, pp. 995-1003.
46. X. Xu, Y. Sun, and Z. Huang, "Defending DDoS attacks using hid- den Markov models and cooperative reinforcement learning," in Proc. Pacific-Asia Workshop Intell. Secur. Inform. Berlin, Germany: Springer, 2007, pp. 196-207.
47. S. Smadi, N. Aslam, and L. Zhang, "Detection of online phishing email using dynamic evolving neural network based on reinforcement learning," Decis. Support Syst., vol. 107, pp. 88-102, Mar. 2018.
48. M. Feng and H. Xu, "Deep reinforecement learning based optimal defense for cyber-physical system in presence of unknown cyber-attack," in Proc. IEEE Symp. Ser. Comput. Intell. (SSCI), Nov. 2017, pp. 1-8.
49. MXNET An Efficient Library for Deep Learning. Accessed: Aug. 13, 2020. [Online]. Available: https://mxnet.apache.org/versions/1.6/

50. S. M. Lee, D. S. Kim, J. H. Kim, and J. S. Park, "Spam detection using feature selection and parameters optimization," in Proc. Int. Conf. Complex, Intell. Softw. Intensive Syst., Feb. 2010, pp. 883-888.
51. N. F. Shah and P. Kumar, "A comparative analysis of various spam clas- sifications," in Progress in Intelligent Computing Techniques: Theory, Practice, and Applications. Singapore: Springer, 2018, pp. 265-271.
52. I. Alsmadi and I. Alhami, "Clustering and classification of email con- tents," J. King Saud Univ. Comput. Inf. Sci., vol. 27, no. 1, pp. 46-57, Jan. 2015.
53. A. A. A. Abdelrahim, A. A. E. Elhadi, H. Ibrahim, and N. Elmisbah, "Feature selection and similarity coefficient based method for email spam filtering," in Proc. Int. Conf. Comput., Electr. Electron. Eng. (ICCEEE), Aug. 2013, pp. 630-633.
54. D. C. Chandrasekar, "Classification techniques using spam filtering email," Int. J. Adv. Res. Comput. Sci., vol. 9, no. 2, pp. 402-410, Apr. 2018.
55. A. A. Elhadi, M. A. Maarof, and A. H. Osman, "Malware detection based on hybrid signature behaviour application programming interface call graph," Amer. J. Appl. Sci., vol. 9, no. 3, p. 283, 2012.
56. D. DeBarr and H. Wechsler, "Spam detection using clustering, random forests, and active learning," in Proc. 6th Conf. Email Anti-Spam. Moun- tain View, CA, USA: Citeseer, 2009, pp. 1-6.
57. A. Ramachandran, N. Feamster, and S. Vempala, "Filtering spam with behavioral blacklisting," in Proc. 14th ACM Conf. Comput. Commun. Secur. (CCS), 2007, pp. 342- 351.
58. A. Sharaff, N. K. Nagwani, and A. Dhadse, "Comparative study of classification algorithms for spam email detection," in Emerging Research in Computing, Information, Communication and Applications. New Delhi, India: Springer, 2016, pp. 237-244.
59. C. Romero, M. Garcia-Valdez, and A. Alanis, "A comparative study of blog comments spam filtering with machine learning techniques," in Soft Computing for Recognition Based on Biometrics. Berlin, Germany: Springer, 2010, pp. 57-72. 344
60. T. Yoshinaka, S. Ishii, T. Fukuhara, H. Masuda, and H. Nakagawa, "A user- oriented splog filtering based on a machine learning," in Recent Trends and Developments in Social Software. Berlin, Germany: pringer, 2008, pp. 88-99. K. Indira and E. C. Joy, "Prevention of spammers and Promoters in Video Social Networks using SVM-KNN," Int. J. Eng. Technol., vol. 6, no. 5, pp. 2024-2030, Oct./Nov. 2014. SolarWinds MSP Mail Assure. Accessed: Feb. 16, 2020. [Online]. Available: https://www.solarwindsmsp.com/products/mail361SpamTitan. Accessed: Feb. 16, 2020. [Online]. Available: https://www.spamtitan.com/
61. SPAMfighter. Accessed: Feb. 16, 2020. [Online]. Available: https://www.spamfighter.com/SPAMfighter/Product_Info.aspZEROSPAM. Accessed: Feb. 16, 2020. [Online]. Available: https://www.zerospam.ca/en/home/
62. S. J. Delany, M. Buckley, and D. Greene, "SMS spam filtering: Methods and data," Expert Syst. Appl., vol. 39, no. 10, pp. 9899-9908, Aug. 2012.
63. I. Ahmed, D. Guan, and T. C. Chung, "SMS classification based on naive Bayes classifier and Apriori algorithm frequent itemset," Int. J. Mach. Learn. Comput., vol. 4, no. 2, p. 183, 2014.
64. B. Biggio, G. Fumera, I. Pillai, and F. Roli, "Image spam filtering using visual information," in Proc. 14th Int. Conf. Image Anal. Process. (ICIAP), Sep. 2007, pp. 105- 110.
65. H. Li, X. Xu, C. Liu, T. Ren, K. Wu, X. Cao, W. Zhang, Y. Yu, and D. Song, "A machine learning approach to prevent malicious calls over telephony networks," in Proc. IEEE Symp. Secur. Privacy (SP), May 2018, pp. 53-69.

66. R. Puzis, M. D. Klippel, Y. Elovici, and S. Dolev, "Optimization of NIDS placement for protection of intercommunicating critical infrastructures," in Proc. Eur. Conf. Intell. Secur. Inform. Berlin, Germany: Springer, 2008,pp. 191-203.
67. N. B. Amor, S. Benferhat, and Z. Elouedi, "Naive bayes vs decision trees in intrusion detection systems," in Proc. ACM Symp. Appl. Comput. (SAC), 2004, pp. 420-424.
68. Y. Bouzida and F. Cuppens, "Neural networks vs. decision trees for intru- sion detection," in Proc. IEEE/IST Workshop Monitor., Attack Detection Mitigation (MonAM), vol. 28. Citeseer, 2006, p. 29.
69. D. S. Kim and J. S. Park, "Network-based intrusion detection with support vector machines," in Proc. Int. Conf. Inf. Netw. Berlin, Germany: Springer, 2003, pp. 747-756.
70. B. Sezari, D. P. F. Moller, and A. Deutschmann, "Anomaly-based network intrusion detection model using deep learning in airports," in Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun., 12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE), Aug. 2018, pp. 1725-1729.
71. M. Aljanabi, M. A. Ismail, and V. Mezhuyev, "Improved TLBO-JAYA algorithm for subset feature selection and parameter optimisation in intrusion detection system," Complexity, vol. 2020, pp. 1-18, May 2020.
72. McAfee Network Security Platform. Accessed: Feb. 16, 2020. [Online]. Available: https://www.mcafee.com/enterprise/en-au/products/network-security-platform.html
73. Hillstone S-Series Network Intrusion Prevention System (NIPS). Accessed: Feb. 16, 2020. [Online]. Available: https://www.hillstonenet.com/products/network-intrusion-prevention-system-s-series/
74. NIP2000/5000 Intrusion Prevention System. Accessed: Feb. 16, 2020.[Online]. Available: https://e.huawei.com/en/related-page/products/enterprise-network/security/application-gateway/nip-ips/security_nip2000_5000_ips_v2_en
75. Palo Alto Networks Completes Acquisition of The Crypsis Group. Accessed: Oct. 10, 2020. [Online]. Available: https://www.paloaltonetworks.com/
76. THE SAGAN LOG ANALYSIS ENGINE. Accessed: Feb. 16, 2020. [Online].Available: https://quadrantsec.com/sagan_log_analysis_engine/
77. A. Govindaraju, "Exhaustive statistical analysis for detection of meta- morphic malware," Master's Projects, 2010, p. 66. [Online]. Available: https://scholarworks.sjsu.edu/etd_projects/66/, https://doi.org/10.31979/etd.ucv9-qd8t.
78. S. B. Mehdi, A. K. Tanwani, and M. Farooq, "IMAD: In-execution malware analysis and detection," in Proc. 11th Annu. Conf. Genetic Evol. Comput. (GECCO), 2009, pp. 1553-1560.
79. Malwarebytes. Accessed: Feb. 16, 2020. [Online]. Available: https://www.malwarebytes.com/
80. Norton Power Eraser. Accessed: Feb. 16, 2020. [Online]. Available: https://us.norton.com/support/tools/npe.htmlAVG. Accessed: Feb. 16, 2020. [Online]. Available: https://www.avg.com/en-ww/homepage#pc
81. Bitdefender Antivirus. Accessed: Feb. 16, 2020. [Online]. Available: https://www.bitdefender.com/
82. Q. Su, J. Tian, X. Chen, and X. Yang, "A fingerprint authentication system based on mobile phone," in Proc. Int. Conf. Audio Video-Based Biometric Person Authentication. Berlin, Germany: Springer, 2005, pp. 151-159.
83. Martinelli, F. Mercaldo, A. Saracino, and C. A. Visaggio, "I find your behavior disturbing: Static and dynamic app behavioral analysis for detection of Android

malware," in Proc. 14th Annu. Conf. Privacy, Secur.Trust (PST), 2016, pp. 129-136.c.
84. A. De Paola, S. Gaglio, G. L. Re, and M. Morana, "A hybrid system for malware detection on big data," in Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS), Apr. 2018, pp. 45-50.
85. P. Wang and Y.-S. Wang, "Malware behavioural detection and vaccine development by using a support vector model classifier," J. Comput. Syst. Sci., vol. 81, no. 6, pp. 1012- 1026, 2015.
86. C. Chen, J. Zhang, Y. Xie, Y. Xiang, W. Zhou, M. M. Hassan, A. AlElaiwi, and M. Alrubaian, "A performance evaluation of machine learning-based streaming spam tweets detection," IEEE Trans. Comput. Social Syst., vol. 2, no. 3, pp. 65-76, 2015.
87. What Are STIX/TAXII. Accessed: Nov. 16, 2020. [Online]. Available: https://www.anomali.com/resources/what-are-stix-taxii
88. P. Kumpulainen and K. Hätönen, "Anomaly detection algorithm test bench for mobile network management," Tampere Univ. Technol., Tampere, Finland, 2008. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.618.5065&rep=rep1&type=pdf
89. S. B. Mehdi, A. K. Tanwani, and M. Farooq, "IMAD: In-execution malware analysis and detection," in Proc. 11th Annu. Conf. Genetic Evol. Comput. (GECCO), 2009, pp. 1553-1560.
90. P. Wang and Y.-S. Wang, "Malware behavioural detection and vaccine development by using a support vector model classifier," J. Comput. Syst. Sci., vol. 81, no. 6, pp. 1012- 1026, Sep. 2015.
91. B. Sanz, I. Santos, C. Laorden, X. Ugarte-Pedrero, and P. G. Bringas, "on the automatic categorisation of Android applications," in Proc. IEEE Consum. Commun. Netw. Conf. (CCNC), Jan. 2012, pp. 149-153.
92. M. Odusami, O. Abayomi-Alli, S. Misra, O. Shobayo, R. Damasevicius, and R. Maskeliunas, "Android malware detection: A survey," in Proc. Int. Conf. Appl. Informat. Cham, Switzerland: Springer, 2018, pp. 255-266.
93. I. Santos, F. Brezo, X. Ugarte-Pedrero, and P. G. Bringas, "Opcode sequences as representation of executables for data-mining-based unknown malware detection," Inf. Sci., vol. 231, pp. 64-82, 2013.
94. H.-S. Ham and M.-J. Choi, "Analysis of Android malware detection performance using machine learning classifiers," in Proc. Int. Conf. ICT Converg. (ICTC), Oct. 2013, pp. 490-495.
95. Kamran Shaukat, Suhuai Luo, Vijay Varadharajan, Ibrahim A. Hameed, Min Xu, " A Survey on Machine Learning Techniques for Cyber Security in the Last Decade", IEEE Access, vol. 8, pp. 222310 – 222354, 2020.