IJNRD.ORG ISSN : 2456-4184

INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Comprehensive Assessments of Cybercrime Detection Techniques

Dr. Nagendra Kumar Sahu
HOD
Department of Computer Science
Aadarsh Mahavidyalya Datrenga Raipur (C.G.)

Sadhana Sahu
Asst. Professor
Department of Computer Science
Aadarsh Mahavidyalya Datrenga Raipur (C.G.)

ABSTRACT

Cybercrimes are cases of indictable offences and misdemeanors that involve computers or commutilation tools as targets and commission instruments or are associated with the prevalence of computer technology. Common forms of cybercrimes are child pornography, cyber stalking, identity theft, cyber laundering, credit card theft, cyber terrorism, drug sale, data leakage, sexually explicit content, phishing, and other forms of cyber hacking. They mostly lead to a privacy breach, security violation, business loss, financial fraud, or damage in public and government properties. Thus, this study intensively reviews cybercrime detection and prevention techniques. It first explores the different types of cybercrimes and discusses their threats against privacy and security in computer systems. Then, it describes the strategies that cybercriminals may utilize in committing these crimes against individuals, organizations, and societies. It also reviews the existing techniques of cybercrime detection and prevention. It objectively discusses the strengths and critically analyzes the vulnerabilities of each technique. Finally, it provides recommendations for the development of a cybercrime detection model that can detect cybercrimes effectively compared with the existing techniques.

Keywords: - Security, cybercrime detection techniques, neural network, fuzzy logic, machine learning, and data mining.

I.INTRODUCTION

Cybercrime is defined as any crime conducted using computers or other communication tools to cause fear and anxiety to people or damage, harm, and destroy properties. Cybercrimes have two categories, namely, computerassisted and computer-focused cybercrimes. Examples of computer-assisted cybercrimes are child pornography, fraud, money laundering, and cyber stalking, whereas examples of computer-focused cybercrimes are hacking, phishing, and website defacement [1].

Obtaining correct and official statistics on cybercrimes is challenging because of the culture in which the crimes were committed, the severity of the offences, and the unreported incidents due to the lack of knowledge or societal constraints. Law enforcement plays an important role in these cases because it controls the level of detail that is reported [1]. The first cybercrime incident, in which computer codes were replicated, took place in the 1960s [2]. Many fraud and forgery cases were reported after 1970 when a bank teller at New York's Union Dime Savings Bank embezzled over $1.5 million from customer accounts. A creeper virus was developed by Bob Thomas in 1971 to infect the systems of the Advanced Research Project Agency Network (ARPANET), which was the first network with packet-switching technology and the TCP/IP protocol [2], [3]. In early 1977, an employee at Imperial Chemical Industries stole hundreds of computers and their backups from the company and asked for 275,000 pounds sterling as a ransom [2]. In 1988, Robert T. Morris developed the first computer worm via a computer at the Massachusetts Institute of Technology [4]. In 1994, Russian hackers transferred huge amounts of money from a city bank to bank accounts in Russia, Finland, Israel, Germany, the United States, the Netherlands, and Switzerland [2]. The first phishing attempt was made in 1995 [2]. The Electronic Disturbance Theater was established in 1997, which was responsible for creating electronic versions of site-in tools that are used in protests. Protesters in 1998 used a tool called FloodNet to perform a denial-of-service attack on the website of the president of Mexico. In January 1998, a revenger system operator remotely changed the Supervisory Control and Data Acquisition (SCADA) system of a coal-fired power plant to its emergency mode, and the SCADA system software was then removed; the SCADA system is utilized to control and monitor equipment or a plant in an industrial field d [5]. In 2005, computer systems in a European bank were shut down due to an attack against air conditioning systems, causing the increased temperature in its computer room. In 2006, the Russian Business Network organization was established [2]. This illegal organization conducted many cybercrimes and offered many cybercrime tools and services related to Trojans, spam, and phishing.

It specialized in personal identity theft for resale. In 2011, British intelligence agencies replaced a webpage that described how to make bombs with one that described how to make cupcakes. The literature review of this study covered studies that have been conducted to develop techniques for the detection and prevention of cybercrimes. The existing techniques have been reviewed and analyzed by many review and survey studies. However, the existing review studies either focused on studying certain cybercrimes, such as cyberbullying [6], botnets [7], fake profiles [8], phishing [9], and email spam [10], or reviewing particular detection techniques such as data mining [11], [12], machine learning [13], and deep learning [14]. This study provides a comprehensive review of cybercrime detection techniques, which are categorized based on the use of different detection methods. The study first presents the different types of cybercrimes and discusses their consequences against individuals, organizations, and societies. Second, it comprehensively reviews the existing techniques of cybercrime detection and classifies them into the following categorized techniques: 1) Statistical-based techniques, which focus on analyzing and extracting information from research data to develop effective methods for cybercrime detection; 2) machine learning techniques, which focus on predicting outputs according to a given input data; 3) neural network-based techniques, which are used to find reasonable solutions for cybercrimes; 4) fuzzy logic classifier and genetic algorithm, which intends to minimize possible false alerts that rise during the detection of cybercrimes; and 5) data-mining-based techniques, which are developed to detect cybercrimes using apriori algorithm. Third, this study also covers other techniques that have been developed to detect cybercrimes based on other detection methods, such as computer vision, biometric, cryptography, and forensic tools. Fourth, this study critically analyzes the strengths and drawbacks to evaluate the detection efficiency of the reviewed techniques in terms of accuracy, response time, and falsealarm rates. Lastly, the study provides some recommendations to enhance the efficiency of the existing techniques and increase their detection accuracy. The rest of this paper is organized as follows. Section 2 introduces and defines some types of cybercrimes. Section 3 discusses previous studies on cybercrime detection techniques that use different technologies, such as machine learning and data mining technology. Section 4 discusses datasets. Section 5 presents the conclusions and future work.

II. CYBERCRIME TYPES

Cybercrimes can be divided into several categories [1]. The following subsections name and explain these categories in detail.

A.CYBER TERRORISM

Cyber terrorism is an unlawful action that involves violence against people and properties. It often has political, and racial or ideological purpose. Besides, this type of cybercrimes can spread fear, anxiety, and violence amongst people or sabotage as well as destroy properties (e.g. computers and networks). Cyber terrorism can also affect the availability and integrity of information [2]. Terrorists utilize the Internet for disseminating of propaganda, recruiting individuals, influencing public opinion, and shutting down national infrastructure (e.g., transportation, dams, traffic lights, and energy facilities). An example of cyber terrorism is the Ukrainian attack on a power grid in December 2015, which began with a phishing email. Certain sequences of cyber terrorists create fear and disruption amongst citizens regarding their safety. Such sequences can also influence political decision-making. Serious economic loss, property damage, and violence as a result of cyber terrorism can lead to death and affect the cohesion of society [2].

B. CYBER WARFARE

Cyber warfare is a type of warfare that does not use weapons, but cyberattacks. It can be performed by organizations or groups of hackers without permission from the government, and it can lead to political problems amongst countries [15]. Today, cyberwarfare and cyberattacks are the most common type of warfare. Many cyberwars have taken place in the last 20 years. For example, Russia and Georgia were engaged in a cyberwar in 2008, which have involved several attacks on the Georgian government websites via structured query language (SQL) injection, distributed denial-of-service (DDoS), and cross-site scripting [15]. Both Israel and Arab hackers have committed many cyberwars against each other. For example, in December 2008, Israel attacked a Hamas TV station, Al-Aqsa, to broadcast a cartoon movie of Hamas' leader being killed, which was tagged with Arabic comments that stated, ''Time is running out'' [15]. In 2007, a group of hackers hacked several Estonian government websites. The Estonian government blamed Russia for these attacks. In Ukraine, on December 23, 2015, electrical power was disconnected all over the country. Three regional electrical power distribution companies, called oblenergos, and more than 50 substations were affected by malicious attacks and went offline [16]. Approximately 225,000 customers were affected for a few hours. All customers were unable to contact the center via the phone to report electricity outages due to the attack. Power was manually brought back after six hours. Malware was found in three different companies in different infrastructure sectors, but their operations were not affected [16]. Another attack on a Ukrainian power station occurred one year later, cutting electricity to certain ministries and the national railway system [17]. All the affected oblenergos proceeded to work under restricted conditions and manually attempted to recover after the attack. However, the attackers implemented techniques to slow down and stop the recovery process [18]. One such technique is remote disconnection of the uninterruptable power supply system [19]. The attackers also have changed the passwords of legitimate users. Therefore, they were not able to log-in to the system during the recovery process. It took the power stations off for six months to recover from the attack. The attackers replaced legitimate firmware with malicious firmware, which destroyed gateways and caused them to be unrecoverable. Thus, the decision maker of the power stations had to buy new devices and integrate them into the system, but this has involved a very high cost [20].

C. CYBER ESPIONAGE

Espionage refers to any action that involves spies and the theft of important and sensitive information for the benefit of rival companies or foreign governments. Cyber espionage uses computers to conduct missions [15]. In December 2007, approximately 300 British companies suffered from cyber espionage attacks by Chinese organizations [15]. In addition, many organized attacks were made on the computers and networks of the US Department of Defense from 2003 to 2006 by China. These organized series of attacks were called ‘‘Titan Rain.’’

D. CHILD PORNOGRAPHY

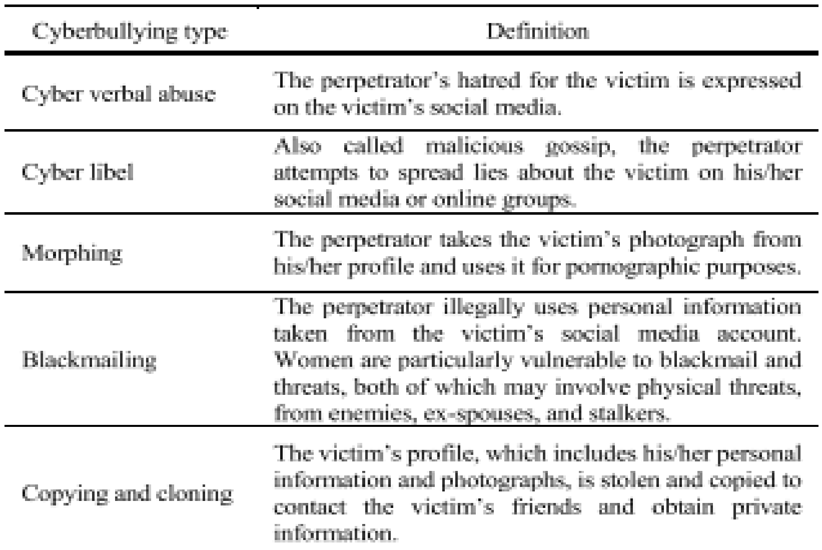
Child pornography refers to pictures, videos, and audio recordings of children wearing inappropriate, few, or no clothes who are in inappropriate positions, specifically sexual positions. Many studies have been conducted to minimize the number of child pornography cases [21]. In general, child pornography contents are distributed for two purposes either for profit or non-profit. For profit purposes, the child pornography are sold in many websites. For non-profit purposes, P2P network can be used to share and distribute those child pornography contents

The law considered any production, possession, or distribution of any type of digital content of child pornography as a serious crime. This is including self-image, trusting others, and disruptions in sexual development. On the other hand, the consequences of this crime on the child side are very harmful and it could last for long time especially the psychological consequences. Those types of consequences and problems will increase if the digital content distributed in the Internet and the child could be a victim for cyber-criminals who are targeting children for sexual purposes.

E. CYBER BULLYING

The increased usage of social media and technology by people of different ages and genders increases the likelihood of unwanted behaviors such as bullying. Bullying is one of the most negative experiences that a person can be faced with, especially during childhood. Most people who experience bullying are children, teenagers, and women. Bullying can inflict emotional and mental harm, and it can affect people’s personality [22]. Victims may receive harmful and rude tweets, messages, or posts that suggest violence, harass the victims, or threaten their lives. Cyber bullying is a type of cybercrime that includes any activity that is harmful to a person, including identity theft, credit card theft, bullying, stalking, and psychological manipulation [22]. Table 1 describes some of the cyber bullying types that could victim go through.

TABLE 1. Cyberbullying types.



Cyberbullying type	Definition
Cyber verbal abuse	The perpetrator's hatred for the victim is expressed on the victim's social media.
Cyber libel	Also called malicious gossip, the perpetrator attempts to spread lies about the victim on his/her social media or online groups.
Morphing	The perpetrator takes the victim's photograph from his/her profile and uses it for pornographic purposes.
Blackmailing	The perpetrator illegally uses personal information taken from the victim's social media account. Women are particularly vulnerable to blackmail and threats, both of which may involve physical threats, from enemies, ex-spouses, and stalkers.
Copying and cloning	The victim's profile, which includes his/her personal information and photographs, is stolen and copied to contact the victim's friends and obtain private information.

After children, women are most vulnerable to cybercrimes because, women tend by nature to be sociable. They easily acquaint themselves with virtual friends or online groups with whom they can discuss cooking techniques, children and family issues as well as post-pregnancy tips. Halder and Karuppannan [22] have suggested that this acquaintanceship can lead to cybercrimes, which highlighting different types of victimization.

F. PHISHING

Phishing is one of the most popular attacks due to its direct connection to the end user. In such cases, the attacker attempts to fool the end user to provide him/her with sensitive information.

Phishing involves a combination of spoofing techniques and social engineering. The victim receives an email asking him/her about sensitive information, warning him/her about an attack, and persuading him/her to install new protection software that is actually malware. Alternatively, a phishing email may contain a link to a fake website [9].

One of the important defensive methods is not to click on a link that appears in a suspicious email. Other ways to protect yourself from phishing attacks are to only visit safe websites that have 'https' in their URL and to install anti-virus software, firewalls, and anti-phishing toolbars [23].

G. DENIAL-OF-SERVICE ATTACK

Denial-of-service (DoS) attacks are a major online threat in which the attacker compromises the availability of services. DoS crashes compromised systems with a huge number of requests, such as Internet Control Message Protocol (ICMP) and SYN floods, causing the systems to get crashed and stop providing the intended service. Another type of DoS attack called a distributed denial-of-service (DDoS) attack, the attacker has access to many channels in a network, and each victim becomes an agent to attack another system, like a zombie [24]. Figure 1 illustrates an example of a DDoS attack. DoS and DDoS attacks take place through the following methods:

1) ICMP FLOOD ATTACK OR SMURF ATTACK

ICMP is a connectionless protocol used to diagnose networks and identify errors. The attacker overwhelms the target server with a huge number of ICMP messages, and the victim server deals with each message and processes it until the server becomes overwhelmed and crashes [25], [26].

2) SYN FLOOD ATTACK

The attacker overwhelms the target system with a flood of SYN attacks to prevent the targeted system from responding to legitimate users [26].

3) TEARDROP ATTACK

The attacker overwhelms the target system with disorganized and overlapped packets. Legitimate senders break messages into organized packets, but the attacker manipulates packets to make them large with large payloads. This causes the target system to become overwhelmed and attempt to reassemble the manipulated and overlapped packets until the system can no longer respond to legitimate users [25]. DDOS attacks can be prevented or mitigated using two methods: the first is to implement DDOS attack prevention services, and the second is to increase the traffic bandwidth of the company's website [23].

H. SQL INJECTION ATTACK

The SQL injection attack is a type of attack in which the attacker compromises databases using some SQL queries. The attacker can look at the database and retrieve its content before altering or deleting the data [27]. One of the best prevention strategies for this type of attack is to set a high standard level of credentials, such as username and password, for all users [23].

I.FUTURISTIC IN CYBER ATTACKS

Futuristic cyber-attacks can target many new and recent technologies and devices, such as WiFi, health care devices, robots, and drones. These new technologies are highly vulnerable to cyber-attacks. WiFi technology is widely used among users and industries; this can jeopardize the security for such users and companies. Some examples of attacks that could affect WiFi users are the man-in-the-middle attack, the key reinstallation attack (KRACK), and the signal jamming attack [23]. In the health care sector, implantable medical devices (IMDs) suffer from security vulnerabilities that can cause harmful consequences to people's health if exploited. IMDs are electronic devices implanted inside the human body to treat or control disease [28]. Examples of IMDs devices include the following: • Implantable cardioverter defibrillators (ICDs) are devices implanted to monitor the heart rate of the patient [28]. Insulin pumps are devices implanted to deliver insulin regularly [28]. Implantable nerve stimulators that are devices to treat chronic pain via sending electrical current in the human body [28]. Robots are also vulnerable to attacks; those targeted include industrial robots and elder care robots. Drones and unmanned aerial vehicles (UAV) are another target for the attackers. UAVs can be hacked since their on-board chips are not encrypted and they are connected to the ground controller through WiFi. Therefore, they are vulnerable to all of the attacks applied on WiFi technology, including man-in-themiddle attacks and signal jamming attacks [23]. Table 2 lists the current cybercrimes and summarizes their features, level of crime, and targets.

II. CYBERCRIME DETECTION TECHNIQUE

The number of cybercrimes has rapidly increased as none of the traditional cybercrime detection systems implemented by forensics researchers can completely stop or mitigate them. This is because the victims or targets of cybercrimes (e.g., people, banks, properties, and governments) differ depending on the motivation for the crime (e.g., money, fame, sex, curiosity), and cybercriminals improve their methods and

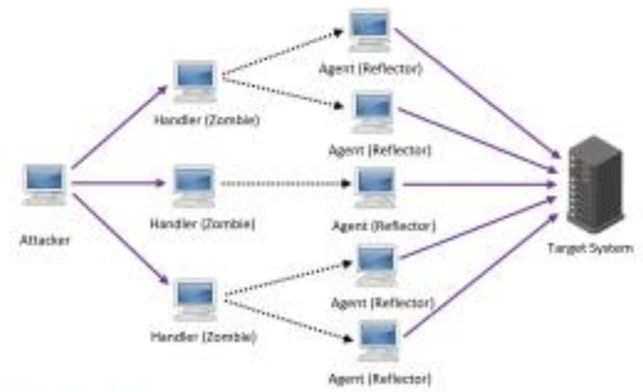


FIGURE 1. An example of a DDoS attack.

Utilize new technologies to commit crimes and achieve their goals. Many prior studies have been conducted to develop methods for detecting cybercrimes. The main categories of these methods are shown in Figure 2 and described in the following subsections.

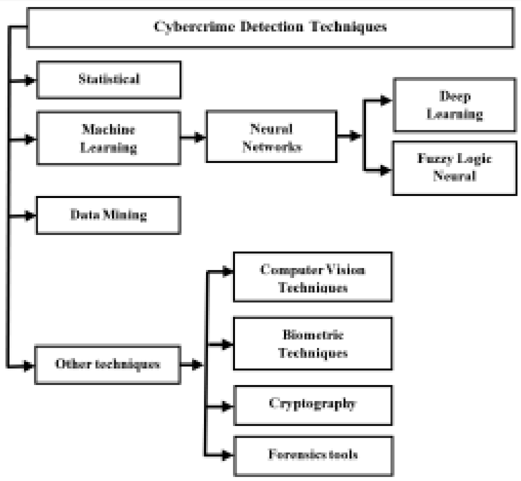


FIGURE 2. Categorization of cybercrime detection techniques.

A. CYBERCRIME DETECTION USING STATISTICAL METHODS

The Hidden Markov Model is one of the best models for detecting cyberattacks. However, it is a time-consuming process. Sultana et al. [29] improved the Hidden Markov Model by minimizing the time required for data training to detect cyberattacks using the N-gram extraction algorithm. This improved Hidden Markov Model utilizes recurrent or repeated patterns in trace files instead of whole trace events. The N-gram extraction algorithm was used during data mining to extract common patterns. As a result, the datatraining time for constructing a system was reduced by 31.96– 48.44%. Liang et al. [30] proposed a filter for an intrusion detection system (IDS) to detect attacks in vehicle ad hoc networks (VANETs) are a special type of networks responsible of monitoring the movement of a group of vehicles without utilizing a base station. It also arranges and manages the communication between the vehicles [31]. This filter was intended to decrease the response time and overhead in the detection process without affecting detection accuracy. The authors utilized the Hidden Markov Model to implement the filter. On the other hand, Qiao et al. [32] proposed an IDS that utilized the Hidden Markov Model and was based on the University of New Mexico (UNM) dataset. Rasmi and Jantan [33] developed a new algorithm for an IDS based on cosine similarity to predict attack intentions. This new algorithm, called the similarity of attack intentions (SAI) algorithm, generates a similarity matrix of previous and known attack intentions that is used to calculate the probability ratio for each attack intention. Similarity is calculated based on the ratio of new attacks to known and predefined attacks. Harrou et al. [34] designed an anomaly detection system to detect TCP SYN flood attacks based on the 1999 DAPRA dataset. TCP SYN floods are utilized in DoS and DDoS attacks. The researchers used the CRPA measure because of its sensitivity to any changes in common patterns of packet flow. They merged the CRPA measure with two statistical methods— Exponentially Weighted Moving Average (EWMA) and Shewhart—to identify the best anomaly detection system. The researchers compared the performance of four mechanisms: EWMA, Shewhart, CPRA-EWMA, and CPRA-Shewhart. The experiment showed that merging the CPRA with the EWMA and Shewhart achieved superior results.

The CPRA-Shewhart mechanism detected attacks with many false alarms, while the CPRA-EWMA mechanism detected attacks without false alarms. Therefore, the CPRA-EWMA mechanism outperformed the CPRAShewhart, EWMA, and Shewhart mechanisms. Abouzakhar et al. [35] developed a system to detect network cybercrimes using a Bayesian learning network approach. The authors have applied their proposed system to a DARPA 2000 dataset of DDoS attacks generated by the Massachusetts Institute of Technology (MIT) Lincoln Laboratory. They evaluated the results using the Life Chart Method. However, it should be noted that the Bayesian network relies on probabilistic models, which are only work well in noisy environments but are unsuitable in real environments. Additionally, this approach is not as feasible as deterministic correlation methods in reality. Wang et al. [36] detected and mitigated a new type of DDoS attack called a link-flooding attack (LFA). LFA attacks can cut off service in very critical areas of a network by flooding them with legitimate low-speed flows. Therefore, normal IDSs, such as an anomaly detection system or signaturebased detection system, cannot detect this type of attack. The researchers proposed a new defense system called LFADefender. While a traditional IDS is installed in a fixed location in a network, LFADefender is adjustable and can change its location in the network in real time. LFA attackers attack the target with high-flow-density links. Therefore, the first task of LFADefender is to find high-flow-density links in the network through software defined networking (SDN) [37], [38]. Software defined networking (SDN) is an architecture that abstracts a control plane from data to achieve more flexibility in network management [39]. After high-density or congested links are detected, rerouting is initiated to avoid the congested links and mitigate—but not stop—the LFA attack. The link density or congestion is monitored by sFlow traffic analyzer software [40]. To stop LFA attacks, the researchers proposed a malicious traffic blocking approach to identify the bot and stop it from affecting the network. This approach monitors and traces the traffic in the network. After rerouting, the attacker will update his or her link map, which contains the target links. If the flow packets appear in the new links again, then they are identified as bot flow packets and the source IP address is identified. Finally, a block flow message will be sent from the SDN controller to block those packets from the network. The traced packet is then utilized to define the bot packets using statistical methods, including calculation of the variance and average of packet numbers and an outlier detection algorithm called, the local outlier factor, to specify the time that the packets suddenly increased in the network. To evaluate this framework, the researchers implemented a test bed using CloudLab, an open platform used to simulate attacks and implement new systems. Birkinshaw et al. [41] implemented an IDS using softwaredefined networking (SDN). The authors have targeted two types of attacks: DoS and port scanning, for which they implemented Credit-Based Threshold Random Walk (CBTRW) and Rate Limiting (RL). A CB-TRW algorithm detects worm infection on a host, whereas an RL algorithm is used to prevent

DoS attacks and to detect the number of requests sent and received from a network interface controller (NIC) [42], [43]. Table 3 summarizes the statistics-based methods of cybercrime detection developed in prior studies.

B. CYBERCRIME DETECTION USING MACHINE LEARNING

Machine learning is the science of predicting outputs based on given input data, also called training data. The machine (i.e., computer) learns how to predict correct and appropriate outputs for specific inputs using the training data. This learning process can be supervised or unsupervised. In the supervised learning method, the training data contain pairs: an input and its corresponding output. The outputs are called labeled outputs because the correct output is already known. The machine tries to learn how pairs are built in order to make its own predictions later. In unsupervised learning methods, the outputs are unlabeled. Therefore, the machine does not know the correct output for each given input. This makes the learning process difficult [44].

One of the basic learning models is the decision tree, which is a classic form of decision-making that is similar to the divide-and-conquer method. There are two types of decision trees: binary and multi-class classification. In a binary classification tree, the response is either ‘‘yes’’ or ‘‘no.’’ The question that is asked is called a ‘‘feature,’’ the response to the question is called the ‘‘feature value,’’ and the rating is called a ‘‘label.’’ Preference for one response over the other is called inductive bias [45]. Researchers have utilized different algorithms from the supervised learning algorithm category, including naïve Bayes and the K-nearest neighbor (KNN), and the unsupervised learning algorithm category, such as K-means, to detect cybercrimes. Several algorithms have been tried to achieve high accuracy and good performance. Some of these studies are presented below.

d-answer website, Nandhini and Sheeba [46] have proposed a cyberbullying detection tool using the Levenshtein algorithm and naïve Bayes classifier. While, Reynolds et al. [47] used a C4.5 decision tree learner and instance-based learner. Both learners identified true positive results with an accuracy of 78.5%. On the other attempt, as a way to detect cyberbullying in YouTube comments, Dinakar et al. [48] used three supervised machine learning methods: JRip, J48, and a support vector machine (SVM). The authors also compared a binary classifier and multi-classifier. In contrast, Al-garadi et al. [49] proposed a tool to detect cyberbullying in tweets. They extracted different types of features from each tweet to be utilized in the classifier to detect cyberbullying. Several classifiers— namely, the support vector machine, naïve Bayes, KNN, and decision tree—were tested to determine the best classifier. The authors concluded that naïve Bayes shows the best performance and has sufficient strength [49].

Uzel et al. [50] utilized text classification to identify cyber terror and extremism (CTE). The researchers assigned numerical weights to terms in order to detect vocabulary related to CTE in texts. The document was converted to a vector. The researchers utilized four weighting methods— namely, term frequency-based, binary, term frequency, and inverse document frequency-based weighting—to computerize the vector. A fuzzy set based on the weighting methods was proposed and implemented. The researchers used SVM and a naïve Bayes multinomial as classifiers to detect CTE. They have also used the antisocial behavior data set in their experiment. The results showed that the fuzzy set-based weighting method with SVM outperformed the other methods, with accuracy of up to 99%. To date, most works have examined cyberbullying only in English-language texts. Only Haider et al. [51] focused on the Arabic language. The researchers used Waikato Environment for Knowledge Analysis (WEKA) because it supports the Arabic language, and they utilized naïve Bayes and SVM to classify texts as either cyberbullying or not cyberbullying. Benferhat et al. [52] proposed a naïve Bayes approach to observe alert correlations and detect cyberattacks as soon as possible before the attack occurs by observing the attack plan. An attack plan is a series of procedures that an attacker follows until he or she achieves the goal. The proposed system detects the attack plan by using the available history of observations. Using the DAPRA 2000 data set, the authors have found that their system reduces false reporting of attacks and does not require an attack scenario or knowledgeable expert to use.

Naïve Bayes is a simple form of a general Bayesian learning network. Therefore, it has the same problem of being probabilistic. It is called ‘‘naïve’’ because it assumes that the variables are independent of each other, which is not correct in reality [53]. Hee et al. [54] implemented a system to automatically detect signals of cyberbullying content in social media texts. They contributed to the field by developing a system to detect cyberbullying with not only aggressive language, but also implicit content, which they explained as difficult as many types of implicit cyberbullying; such as curses, defamation, and encouragement, that may have different types of attitudes. To do so, they utilized binary and linear support vector machine classifiers. They applied the proposed system to texts in English and Dutch, working with a dataset of 113,698 English and 78,378 Dutch ASKfm posts. An SVM classifier was implemented using the LIBLINEAR library in Python due to its high ability to perform large linear classification. After optimization, the new model achieved maximum F1-scores of 58.72% and 64.32% for Dutch and English, respectively.

Vijayanand et al. [55] proposed a new IDS for securing a wireless mesh network using a genetic algorithm for feature selection and SVM as a classifier. The proposed system was tested using a simulated wireless mesh network dataset in Network Simulator 3 (NS3). They achieved high accuracy of attack detection (95.5%). Ofoghi et al. [56] proposed a tool with hybrid features that detects phishing emails by extracting feature vectors. This tool uses four processes: feature vector generation, machine learning, method selection, and inductor and feature evaluation. As another attempt, Zulkefli et al. [57] investigated the methods for making advanced persistent threat (APT) attacks on smartphones. APT attacks are planned attacks combining social engineering and malware; one of the most popular types of APT attacks is phishing. The authors have utilized a decision tree classifier to distinguish legitimate websites from fake websites, achieving accuracy of 90%.

As most IDSs can prevent known pattern attacks, Ahn et al. [58] proposed a new paradigm system to predict unknown attacks. The authors focused on APT attacks, which are more dangerous than normal attacks because the attacker monitors the victim to collect information, identify vulnerabilities, and search for the most privileged users, such as the administrator. Their paradigm system was based on big data techniques, which are used in fields such as machine learning, data mining, and artificial intelligence. The techniques applied by the researchers included prediction using regression analysis, classification using SVM or logistic regression analysis, the relation rule for discovering hidden relationships amongst data, and atypical data mining for analyzing the data that cannot be represented with numbers (e.g., text, images and videos). Darus et al. [59] focused on the Android platform; the popularity of the Android operating system in recent years has encouraged criminals to target it with many types of malware intended to steal sensitive information from users' smartphones. The authors utilized visualization techniques to detect new types of malware by converting APK files to 8-bit greyscale images. A GIST descriptor was used to extract features from the converted images. A GIST descriptor is a holistic filter for an image, it provides a low dimensional image with some information to understand the view in an image [60]. Three types of classification algorithms—KNN, decision tree, and random forest (RF)—were utilized. The authors discovered that RF has better accuracy than the KNN and decision tree methods. In the KNN algorithm, all features in the dataset are equally important and are used in same amounts; thus, no features are labeled as important or more relevant, which is not helpful for detecting cybercrimes with many useless features [45]. Generating images was difficult; half of the malware samples were not converted to images as the APK files were corrupted or did not have the ‘‘.dex’’ class, which is necessary for conversion.

Vuong et al. [61] proposed a method to detect cyberattacks on mobility devices, such as robots, that considers the devices’ mobile nature and energy consumption as well as the physical impact of the attack. Decision tree C 5.0 algorithm is used in this study to implement the classification process. The proposed method on mobile robotic vehicles faces four types of attacks: DoS, SQL injection, and two types of malware (one targeting the network and one targeting the central processing unit) [61]. Al-diabat [62] investigated phishing attacks and ways to minimize this problem. As every phishing attempt is linked to a fake website, Al-diabat tried to detect fake websites by analyzing the features that distinguish between legal and illegal websites, including a lengthy URL, IP address, and an ‘‘@’’ symbol within the URL. The author tested the possibility of reducing the number of website features through feature selection, which filters out the training data to identify specific attributes that best represent the training data and all attributes. The most effective attributes are selected to minimize computational time and resources, reduce the search space by omitting irrelevant features, and ease the classification process. Al-diabat used information gain and symmetrical uncertainty. This type of selection method should not affect the detection of illegal websites. After the most relevant features are selected, the classification process is initiated to test the efficiency of the selected features. The researcher used the C4.5 algorithm, which is a tree-based algorithm, and the incremental reduced error pruning (IREP) algorithm, which is a greedy algorithm. The WEKA software tool was used, and the data were real data obtained from the University of Irvine Repository, Phishtank website, and Yahoo! Directory [62]

Using decision trees to detect cybercrimes has certain drawbacks. For example, detection of cybercrime cannot be applied when the tree is full of leaves, because detailed questions were asked during the investigation process due to the type of crime or incomplete information about the cybercrime or cybercriminal. All machine-learning algorithms are generally affected by noise in the training data, which may be observed at the feature or label level. Table 4 summarizes the machine-learning-based techniques of cybercrime detection. Nath [63] used a clustering algorithm with K-means clustering for data mining to help detect crime patterns. Clusters (of crime) have a special meaning, referring to a geographical group of crimes (i.e., a lot of crimes in a given geographical region). Additionally, the K-means clustering algorithm is sensitive to outliers and noise in data. K-means methods could converge data quickly, but they would not guarantee that when the data get converged would achieve the correct answer. Furthermore, K-means is an unsupervised learning algorithm, and therefore, the correct answers are not known [45].

C. CYBERCRIME DETECTION USING NEURAL NETWORK

A neural network is a simulation of how the human brain works. The brain consists of nerve cells that can learn, which are represented by neurons in the neural network. These neurons can do training and learn by themselves based on previous knowledge. This allows the neural network to find a reasonable solution for similar problems of a similar class for which it is not explicitly trained. Neural networks have a high degree of fault tolerance against noisy input data which is considered an advantage in comparison to machine learning algorithms [64]. Raiyn [65] described some types of cyberattacks as well as some of the strategies that have been used to detect cybercrimes, such as embedded programming, agent-based methods, software engineering, and artificial intelligence approaches. The researcher discussed the detection of cybercrimes in the cloud, presenting some studies that have been done on this topic, and introduced the concept of utilizing IP addresses to determine users’ geographical location (i.e., country, city, and street) as well as for real-time cyberattack detection.

D. CYBERCRIME DETECTION USING FUZZY LOGIC NEURAL NETWORK

Fuzzy logic is a combination of classical and fuzzy sets. It measures the degree of truth, or the degree to which we can say that an item belongs to the set. It does not categorize items into 0 and 1, where 0 indicates that a lack of belonging to a set and 1 indicates belonging to a set. Rather, in fuzzy logic, 0 and 1 indicate extreme cases of truth [83]. This logic is needed for detecting cybercrimes because of the uncertainty and doubt related to collecting evidence. Flexibility is required to assign items to the appropriate group and thus identify the case as a cybercrime or not and the perpetrator as a cybercriminal or not. Fatima et al. [84] defined the soft computer application technique, which is used when a solution cannot be predicted due to lack of supportive and detailed information. Soft computer techniques help deal with and adapt to uncertainty in emotional and physical characteristics. The researchers focused on two soft computing applications: neuro-fuzzy logic and ANN. They compared the two soft computing applications, and the results showed that neuro-fuzzy logic is superior for detecting cybercrimes. Ahmed and Mohammed [85] have utilized the fuzzy minmax approach to detect the attackers' intentions in real time. The process involved two steps. In the first step, the pattern of the attack is determined. While in the second step, the intention of the attack is identified by investigating the similarities between the characteristics of the pattern and the evidence that was collected from the attack by utilizing a fuzzy minmax neural network. Chandrashekhar and Kumar [86] proposed an IDS using a fuzzy min-max neural network and tested it with the KDD '99 data set. In contrast, Aldubai et al. [87] proposed an IDS to detect cybercrimes utilizing a fuzzy min-max neural network classifier and Principal component analysis (PCA) as a feature extraction algorithm. They tested this system using KDD '99 and NSL-KDD. Azad and Jha [88] proposed a new IDS utilizing a fuzzy min-max neural network as a classifier and a genetic algorithm to optimize the hyberbox. This IDS was tested using KDD '99. A year later, Azad and Jha [89] proposed another IDS utilizing a fuzzy min-max neural network as a classifier and particle swarm for optimization, again testing it with KDD '99. Shalaginov et al. [90] emphasized the importance of utilizing soft computing applications in forensics investigations due to the large amount of data that must be analyzed to identify evidence to help investigators. In normal methods, this process consumes time and resources. Soft computing applications, such as fuzzy logic, machine learning and data mining, facilitate big data analytics to assist investigators in detecting cybercrimes and criminals. On the other hand, Barraclough et al. [91] utilized fuzzy logic to detect phishing attacks using five different tables in which 288 features were stored with two-fold cross validation. They achieved high accuracy. Saidi et al. [92] aimed to identify cyberterrorist committees amongst other committees. They used an evidential C-means (ECM) algorithm to cluster network data from the John Jay ARTIS Transnational Terrorism (JJATT) database and Global Terrorism Database (GTD). The researchers tried to improve Constrained Evidential C-Means (CECM) clustering process using two constraints: must-link and cannot-link. Must-link means that two objects must be classified in the same cluster, while cannot-link means that two objects cannot be allocated to the same cluster. After these constraints were applied, a new algorithm, called the constrained ECM algorithm, was proposed.

E. CYBERCRIME DETECTION USING DATA MINING

Sindhu and Meshram [93] have proposed a system for detecting cybercrimes that uses an a priori (i.e., data mining) algorithm. The researchers started from case reports, extracting and determining the attributes/variables of the cases. The a priori algorithm was applied to the set of variables to identify frequent item sets. Although the proposed algorithm was not implemented, the algorithm is useful for detecting the attributes and variables of cybercrime case reports. The researchers utilized visualizations, such as bar chart or graphs, to make analysis easier for investigators Shahresani et al. [94] proposed a new system called a visual threat monitor, which combines data mining and visualization to detect botnet behavior in a network. Data mining is applied to analyze patterns of network packets using packet trace files to distinguish between regular and irregular packets. It can extract adequate data for analysis even when a large amount of data is contained in packet trace files. The authors have used several visualization techniques, such as histograms, grid visualizations, and scatter plots, to help the network administrator detect botnets easily. They have also implemented data mining techniques to achieve accurate results for classification, clustering, aggregation, statistical analysis, and flow correlation. They finally have clarified the differences between the techniques to determine which was able to most accurately detect cybercrimes

F. CYBERCRIME DETECTION USING OTHER TECHNIQUES

This subsection covers other techniques that have been developed to detect cybercrimes based on other detection methods such as computer vision, biometric, cryptography, and forensic tools. Computer vision techniques focus on analyzing and interpreting images [101]. Computer vision techniques have been used to detect cybercrimes, especially phishing, by analyzing the URLs of websites to determine whether they are legitimate or fake. An example of such research was conducted by Rao and Ali [102], who suggested a technique to detect phishing websites by combining a whitelist and visual similarity-based technique. They utilized a speededup robust features (SURF) detection tool to extract features from fake and phished websites. The whitelist, which contains all legitimate URLs, was used to check URLs. Then, a visual similarity-based technique was used to identify the legitimacy of URL via finding the most similar scores either it is legitimate or suspicious URLs. Another researchers used biometric techniques to defend cyber crimes such as Ahmed et al. in [103] proposed an approach to be applied in Bangladesh to detect cybercrimes over the Internet. The new framework requires each Internet's user to register a national ID and password to gain access to the Internet, and foreigners can gain access using their visa's number. Then, the users' faces and fingerprints are scanned and saved into the cloud for biometric verification. Next, users must provide their birth certificate number. Finally, either a phone number or email address is required to complete the activation process. This process would ensure that only legitimate users could gain access to the Internet. The Bangladesh Telecommunication Regulatory Commission will verify users' Internet ID and password in the cloud, and users will gain access to the Internet. All of their activities will be saved in an activity log in the cloud to detect potential cybercrimes. The proposed architecture was tested on 16 volunteers using a network simulator called Packet Tracer. The results showed that the proposed framework could accurately detect cybercrimes. Cryptography is another methodology that has been utilized to detect cyber crimes, where Derhab et al. [104] tackled the spam botnet detection problem via presenting a security framework called Spam Trapping System (STS) which is responsible for providing a third line of detecting and preventing the spam botnet from spreading to the other hosts. Spam Trapping System uses encrypted emails to distinguish

between the legitimate emails and spam emails. This distinguish process uses cryptographic key in legitimate email. The users, the email application, and STS system know the cryptographic key. On the other hand, spam emails are not encrypted with known key. Therefore, those spam emails are not sent outside the host. By this procedure, the third line of protection is created and the spam email is prohibited from going outside the host.

IV.CYBERCRIME TESTING DATASETS

A review of benchmark datasets was presented in [106]. The KDD '99 data set was generated in 1999 by Stolfo et al. [107]. This dataset focuses on four types of attacks: DoS, U2R, remote to local, and probing attacks. However, Abubakar et al. [106] mentioned that the KDD '99 dataset is no longer efficient for IDSs due to the fact that it is an old dataset, and there have been many cybercrimes happened within the psat 20 years, hence it will provide inaccurate results. In addition, Tavallaee et al. [108] stated that about 78% and 75% records in the training set and test set, respectively, are duplicates, which will affect the evaluation process for the detection algorithm. Thus, NSL-KDD was created in 2009 by Tavallaee et al. [108]. This dataset consists of KDD dataset records, minus all the duplicate or redundant records in the training and testing data sets. On the other hand, DAPRA 2000, which includes DDoS attacks, was generated in 2000 by the MIT Lincoln Laboratory [109]. While, Abubakar et al. [85] also reviewed the University of New Mexico (UNM) dataset, which was proposed in 2004 [110]. UNM has several limitations, including a limited scope of cybercrimes, a focus on a single process, and an incomplete sampling of the target operating system [111]. Creech and Hu [111] generated a new benchmark dataset called Australian Defence Force Academy Linux (ADFALD12) in 2013. It consists of system call traces and focuses on six types of attacks: Hydra-FTP, Hydra-SSH, Adduser, JavaMeterpreter, Meterpreter, Webshell [112].

Moustafa and Slay [113] presented the UNSW-NB15 dataset, which is network-based. This dataset focuses on nine types of attacks: fuzzers, backdoors, DoS, exploits, reconnaissance, shellcode, worms, analysis (port scan, HTML file penetration, spam), and generic (a technique that works against all block ciphers). The CICIDS2017 dataset was presented in 2017 by the Canadian Institute of Cyber Security [93]. It contains 14 types of attacks.

V. CONCLUSION

The comprehensive review in this paper has covered several types of cybercrimes and analyzed numerous studies regarding their achieved detection rates as well as some of their limitations. The presented state of the arts in this paper has been evaluated and a comparison was carried out via some tabulated information as a way to demonstrate their results to identify their respective advantages and disadvantages. This study has also intensively discussed the available datasets that have been used by previous studies. Finding the proper dataset for testing and evaluating the research's method for cybercrime detection are critical challenges. The unavailability of benchmark datasets is an inevitable consequence of the lack of cooperation between law enforcement and researchers in terms of cybercriminal data collection. Another challenge is the diversity of cybercrimes, as they may happen within different platforms such as Twitter, YouTube, Instagram, or through networks; which involve different types of datasets.

To overcome the availability challenge of cybercrime datasets, it is recommended to create cybercriminal profiling that can be used by the researchers as cybercrime datasets. However, creating cybercriminal profiling requires a serious collaboration between law enforcement and researchers as well as governmental regulators. Since the information that can be included in the cybercriminal profiling, which is mostly critical, sensitive, and private, the legality for revealing this information is questionable. For this reason, researchers should find a method to protect data privacy; by these means, they may benefit from the data of cybercriminals provided by law enforcement for research purposes while also maintaining their privacy

REFERENCES

[1] M. Yar and K. F. Steinmetz, Cybercrime and Society. Newbury Park, CA, USA: Sage, 2019. [2] B. Akhgar, A. Staniforth, and F. Bosco, Cyber Crime and Cyber Terrorism Investigator's Handbook. Rockland, MA, USA: Syngress, 2014.

[3] M. Rouse. (2017). Arpanet. Accessed: Apr. 26, 2020. [Online]. Available: https://searchnetworking.techtarget.com/definition/ARPANET

[4] (2018). The Morris Worm. Accessed: Jan. 28, 2020. [Online]. Available: https://www.fbi.gov/news/stories/morris-worm-30-years-since-firstmajor-attack-on-internet-110218

[5] V. Beal. (Apr. 27, 2020). SCADA—Supervisory Control and Data Acquisition. [Online]. Available: https://www.webopedia.com/TERM/S/ SCADA.html

[6] S. Nadali, M. A. A. Murad, N. M. Sharef, A. Mustapha, and S. Shojaee, ''A review of cyberbullying detection: An overview,'' in Proc. 13th Int. Conf. Intellient Syst. Design Appl., Dec. 2013, pp. 325–330.

[7] A. Karim, R. B. Salleh, M. Shiraz, S. A. A. Shah, I. Awan, and N. B. Anuar, ''Botnet detection techniques: Review, future trends, and issues,'' J. Zhejiang Univ. Sci. C, vol. 15, no. 11, pp. 943–983, Nov. 2014.

[8] D. Ramalingam and V. Chinnaiah, ''Fake profile detection techniques in large-scale online social networks: A comprehensive review,'' Comput. Electr. Eng., vol. 65, pp. 165–177, Jan. 2018.

[9] A. N. Shaikh, A. M. Shabut, and M. A. Hossain, ''A literature review on phishing crime, prevention review and investigation of gaps,'' in Proc. 10th Int. Conf. Softw., Knowl., Inf. Manage. Appl. (SKIMA), Dec. 2016, pp. 9–15.

[10] W. Z. Khan, M. K. Khan, F. T. B. Muhaya, M. Y. Aalsalem, and H.-C. Chao, ''A comprehensive study of email spam botnet detection,'' IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 2271–2295, 4th Quart., 2015.

[11] H. Hassani, X. Huang, E. S. Silva, and M. Ghodsi, ''A review of data mining applications in crime,'' Stat. Anal. Data Mining, ASA Data Sci. J., vol. 9, no. 3, pp. 139–154, Jun. 2016.

[12] M. BinJubier, A. A. Ahmed, M. A. B. Ismail, A. S. Sadiq, and M. K. Khan, ''Comprehensive survey on big data privacy protection,'' IEEE Access, vol. 8, pp. 20067–20079, 2019.

[13] C.-F. Tsai, Y.-F. Hsu, C.-Y. Lin, and W.-Y. Lin, ''Intrusion detection by machine learning: A review,'' Expert Syst. Appl., vol. 36, no. 10, pp. 11994–12000, 2009.

[14] A. Aldweesh, A. Derhab, and A. Z. Emam, ''Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues,'' Knowl.-Based Syst., vol. 189, Feb. 2020, Art. no. 105124.

[15] J. Carr, Inside Cyber Warfare: Mapping the Cyber Underworld. Newton, MA, USA: O'Reilly Media, 2011, p. 316.

[16] (2016). ICS Alert (IR-ALERT-H-16-056-01): Cyber-Attack Against Ukrainian Critical Infrastructure. Accessed: Jan. 9 2019. [Online]. Available: https://www.us-cert.gov/ics/alerts/IR-ALERT-H-16-056-01

[17] (2017). Ukraine Power Cut 'Was Cyber-Attack'. Accessed: Jan. 9 2019. [Online]. Available: https://www.bbc.com/news/technology-38573074

[18] V. Butrimas. (2016). Threat Intelligence Report Cyberattacks Against Ukrainian ICS. Accessed: Sep. 9, 2019. [Online]. Available: https://www. sentryo.net/wp-content/uploads/2017/10/EBOOK-UKRAINIANCYBERATTACKS-OCT-2017.pdf

[19] K. J. Higgins. (2016). Lessons From The Ukraine Electric Grid Hack. Accessed: Sep. 9, 2019. [Online]. Available: https://www.darkreading. com/vulnerabilities—threats/lessons-from-the-ukraine-electric-gridhack/d/d-id/1324743

[20] K. Zetter. (2016). Inside the Cunning, Unprecedented Hack of Ukraine's Power Grid. [Online]. Available: https://www.wired.com/2016/03/insidecunning-unprecedented-hack-ukraines-power-grid/

[21] N. A. Mutawa, J. Bryce, V. N. L. Franqueira, and A. Marrington, ''Behavioural evidence analysis applied to digital forensics: An empirical analysis of child pornography cases using P2P networks,'' in Proc. 10th Int. Conf. Availability, Rel. Secur., Aug. 2015, pp. 293–302.

[22] D. Halder and K. Jaishankar, ''Cyber socializing and victimization of women,'' Temida, vol. 12, no. 3, pp. 5–26, 2009.

[23] S. S. Chakkaravarthy, D. Sangeetha, M. Venkata Rathnam, K. Srinithi, and V. Vaidehi, ''Futuristic cyber-attacks,'' Int. J. Knowl.-based Intell. Eng. Syst., vol. 22, no. 3, pp. 195–204, Nov. 2018.

[24] C. Douligeris and D. N. Serpanos, Network Security: Current Status and Future Directions. Hoboken, NJ, USA: Wiley, 2007.

[25] (Nov. 26, 2019). DoS (Denial of Service) Attack Tutorial: Ping of Death, DDOS. [Online]. Available: https://www.guru99.com/ultimate-guide-todos-attacks.html

[26] H. Dalziel. (Nov. 26, 2019). 5 Major Types of DOS Attack. [Online]. Available: https://www.concise-courses.com/5-major-types-of-dosattack/

[27] (May 6, 2020). SQL Injection. [Online]. Available: https://portswigger. net/web-security/sql-injection

[28] A. Tabasum, Z. Safi, W. AlKhater, and A. Shikfa, ‘‘Cybersecurity issues in implanted medical devices,’’ in Proc. Int. Conf. Comput. Appl. (ICCA), Aug. 2018, pp. 1–9.

[29] A. Sultana, A. Hamou-Lhadj, and M. Couture, ‘‘An improved hidden Markov model for anomaly detection using frequent common patterns,’’ in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2012, pp. 1113–1117.

[30] J. Liang, M. Ma, M. Sadiq, and K.-H. Yeung, ‘‘A filter model for intrusion detection system in vehicle ad hoc networks: A hidden Markov methodology,’’ Knowl.-Based Syst., vol. 163, pp. 611–623, Jan. 2019.