State-Of-The-Art Machine Learning and Deep Learning: Evolution of Intelligent Intrusion Detection System against Wireless Network (Wi-Fi) Attacks In Internet of Things (Iot)

Nivaashini.M, Thangaraj.P

Abstract- With the quick technical developments of devices and sensors, Wi-Fi have turned out to be a widespread technology for the Internet of Things (IoT). However, the benefit of wireless networks and IoT comes with a cost, which is mainly due to the concern of security and privacy. The distributed nature, multihop data forwarding, and open wireless medium are the factors that make wireless network highly vulnerable to security attacks at various levels. For more than two decades, Intrusion Detection Systems (IDSs) plays an important role in detecting and preventing such security attacks. Still, applying traditional IDS techniques to wireless network of IoT is difficult due to its particular characteristics such as constrained-resource devices, specific protocol stacks, and standards. Thus, a wide look at relating Intrusion Detection System (IDS) with machine-learning procedures in scholastic world and in business field have been done commonly. Yet, massive information and complications to acquire data occurrences in machine-learning-based IDS are sizzling challenges and correspondingly not sophisticated enough to handle persistently erratic wireless network conditions rising from the incredible network traffic evolution. Therefore, Deep learning, the modern revolution in the machine learning & intelligence zone, seems to be a feasible method in scheming Intellectual IDS. In this paper, a review on IDS research efforts by means of machine learning and deep learning practices in wireless networks of IoT has be offered along with the summary of upcoming research guidelines in IDS using deep learning procedures to overwhelm the limits of earlier typical machine learning based IDSs.

Keywords – Intrusion Detection System, Wireless Networks, Internet of Things, Machine Learning Techniques, Deep Learning Techniques.

I. IoT AND ITS WIRELESS COMPONENTS

The Internet organizes a network system which is assorted and framework in reality. As a matter of fact, more than 3 billion individuals pushed toward the Internet in 2014. In addition, there are a comparative number of versatile interests (6.8 billion) as there are individuals on earth [4]. Generally speaking, in 2014, 2.5 Exabyte have been surveyed as an advantageous framework information development [3]. An Exabyte of twenty-four.3 are evaluated compound yearly and located that it's been ascended systematically at a rate of fifty-seven p.c in 2019 [1]. This will be attributable to numerous ingenious parts together with the expansion of bit screen gadgets (cell 23yj, tablets, and such), and, totally, the evolvement and specific sweetening of remote and useful advances. Then again, the net of Things (IoT) could be a quickly creating heterogeneous course of action of connected sensors and actuators mounted to a good extent of ordinary parts. Fig. 1. demonstrates the quick development of IoT by 2020 [2][5]. In an IoT Network, a part of the critical developments is remote individual zone organize viz. 6LoWPAN, ZigBee, Bluetooth. Besides, on a possibly higher remote system scale, Wi-Fi, remote LAN advancement, will be used and ought to be maintained [8].

A wide degree of utilizations and contraptions have been covered under IoT. Applications of Wi-Fi in the IoT has been appeared in [7]. Today, for the most part every house, working condition, bistro, and school it is spread. Wi-Fi is now a most familiar term while suggesting interfacing with the Internet through a remote way. The expansive acknowledgment of Wi-Fi settles on it a first advancement choice for some IoT applications. By the by, in some IoT applications, the choice of development is obliged to the contraptions gear capacities, low-control use essentials, and the general cost Numerous IoT devices require the usage of an ease and low-control remote advancement while interfacing with the Internet [9]. For the most part, vitality use has reliably been a confining variable in various remote sensor sort out applications. This restricting variable will proceed as a basic test confronting the upgrade of different solicitations in the IoT. To be honest, for the headway of

the IoT, low-control use is a noteworthy need that should be met and besides low-control utilization, other related basics should be considered. For instance, the expense of advancement, security, ease (simple to utilize and oversee), remote data rates and ranges, among others, for instance, those reported in [6], are fundamental necessities that require thought. Many creating remote advances, for instance, ZigBee and Bluetooth are battling to outfit the IoT with a low-control organize plan. Diverse remote advancements, for instance, the IEEE 802.11ah, LoRa, and 6Lowpan traditions are growing too [10]. They offer equivalent low-control remote system availability clarifications for the IoT.

Revised Manuscript Received on 10 January 2019.

Nivaashini M, Department of Computer Science & Engg, Bannari Amman Institute of Technology, Erode, Tamil nadu, India,

Thangaraj P, Department of Computer Science & Engg, Bannari Amman Institute of Technology, Erode,Tamil nadu, India,

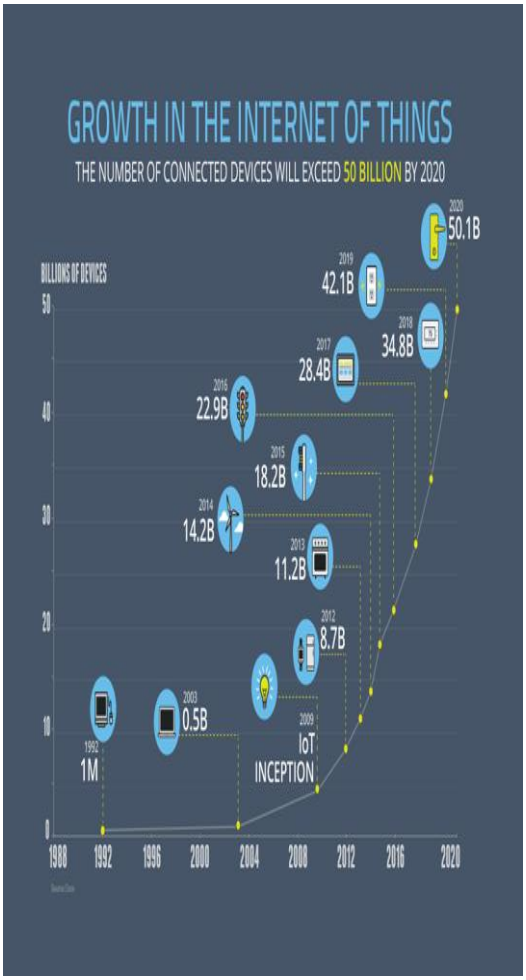

Fig. 1. IoT Growth by 2020

The remainder of the paper is organized as follows. In Section II, provides an overview for the need of different types of IDS in wireless network (Wi-Fi) along with the dataset and attack signatures. In Section III, an overview of various machine learning algorithms and its types are discussed. Several commercially available deep learning platforms are also described in the section. Next, in Section IV, the state-of-the-art machine learning based IDS applications in various wireless networks are extensively surveyed. In Section V, an overview of various deep learning algorithms and its types are presented. The state-of-the-art deep learning-based IDS applications in various wireless networks are deliberated in Section VI. Finally, the paper is concluded in Section VII with the investigation report and future research scope.

II. OVERVIEW OF INTRUSION DETECTION SYSTEM, ITS TYPES AND DATASET

This paper principally centers IoT with IEEE 802.11 remote innovation (Wi-Fi). By 2020 [11] remote system is predicted to speak to 66% of total Internet development with 66% of IP action foreseen that would be made by Wi-Fi and cell devices so to speak. In spite of the fact that remote systems for instance, IEEE 802.11 have been commonly sent to outfit customers with flexibility and versatility as rapid neighborhood, distinctive issues, for instance, assurance and security have raised. The quick spread of Internet of Things (IoT)- engaged devices has brought about remote systems getting to be to both detached and dynamic assaults, the amount of which has grown radically [12]. Cases of these assaults are pantomime, flooding, and infusion assaults. These flare-ups may discharge sensitive data or bother customary undertakings which prompts an enormous cash related setback. The most winning associations influenced by security scenes are budgetary organizations related associations. Pursued by data and correspondences, fabricate, retail and medicinal services [14]. This condition forces to strengthen the wellbeing endeavors in the structures and furthermore in the remote systems Subsequently, Intrusion Detection System (IDS) transforms into a standard wellbeing exertion in system security.

A. Intrusion Detection System

Not at all like firewall, IDS by and large arranged inside the framework to screen each and every inside system traffic movement. In order to make the structure consistent, one may think to have together an IDS and firewall to be connected. It has been depicted that IDS is a mechanization of impedance divulgence process in discovering exercises of infringement of standard security practices or security procedures in frameworks [15]. Other than seeing the security scenes, IDS in addition has particular breaking points: recording existing dangers and hindering foes [15]. IDS requires unequivocal properties which goes about as an idle countermeasure, screens whole or part of frameworks just and focuses high strike recognizable proof rate and low false alert rate. Table I gathers the examination of IDS types.

B. Types of IDS based on the Position

IDSs can be disengaged in perspective of their territory in the framework structure and approach utilized as appeared in Fig.2 and Fig. 3. By methods for arranging the IDS segment in the system framework, IDSs can be seen into 3 types: organize based IDS (NIDS), have based IDS (HIDS) and half and half based IDSs. The NIDS is the prime IDS which helps the framework containing IDS module. As a result, it can screen entire framework structure bargains and has a basic layout of the framework system. Then again, HIDS puts the IDS part on every customer of the framework system. The section can just watch the ingoing or dynamic arrangements of the relating customer provoking focal point impression of the express customer. Two sorts of IDSs have explicit drawbacks– the NIDS may exhaust the stack by then oversights some poisonous happenings, while the HIDS does not have the structure of the entire framework system ideally having less remarkable job that needs to be done over the NIDS. Consequently, the half breed-based IDS put IDS segments in the system framework alongside clients to screen similarly specific clients and system diagram meanwhile.

	Anomaly Dependent	Mis-use Dependent	Specific Dependent
Methodology	Classify rare action patterns	Classify well-known attack patterns	Classify damage of pre-defined rules
Attack Discovery Rate	Less	More	More
False Positive Rate	More	Less	Less
New Attack Discovery	expert	Inexpert	Inexpert

Table I: Types of IDS Vs Detection Methodology

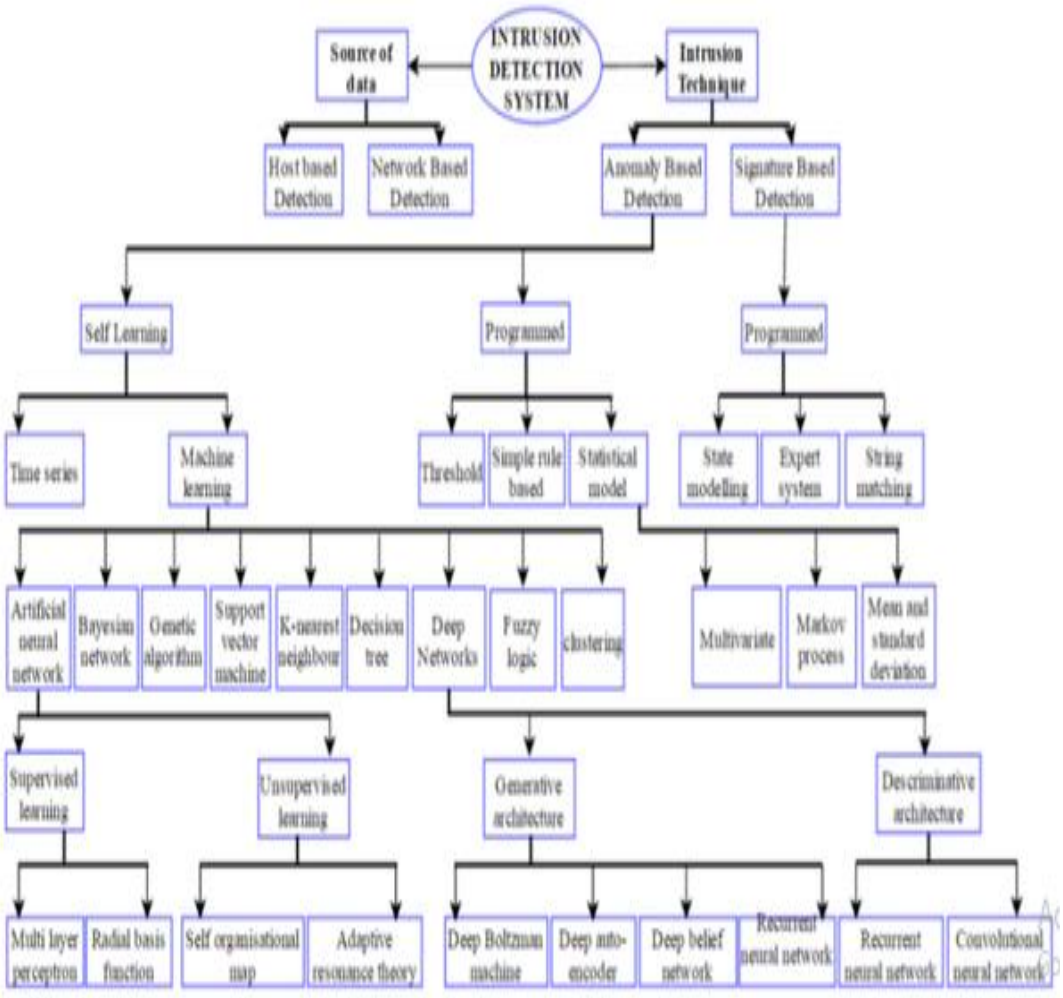

Fig. 2. Overview of IDS types and approaches

C. Types of IDS dependent on the Detection

In the advanced case, in light of the recognizable proof technique, IDSs can be detached into 3 particular makes: abuse based IDS, peculiarity based IDS, and detail based IDSs. An abuse based IDS or mark based IDS [13], searches for some malignant happenings by arranging the remarkable signs or setups of ambushes with the watched framework arrangements and outfits comprehended assault disclosure; regardless, new or cloud strikes are hard to be seen. An idiosyncrasy based IDS separates a strike by outlining typical direct and a while later if there is any deviation, it triggers an alarm. The idea of this IDS is its capacity for obscure assault revelation.

Then again, misuse-based IDS if all else fails accomplish higher disclosure executionif assaults are well known as compared to irregularity-based IDS. Finally, a specific-based IDS physically depicts an arrangement of standards and hindrances to express the typical activities [16]. Table I condenses the examination of IDS types subject to the logic.

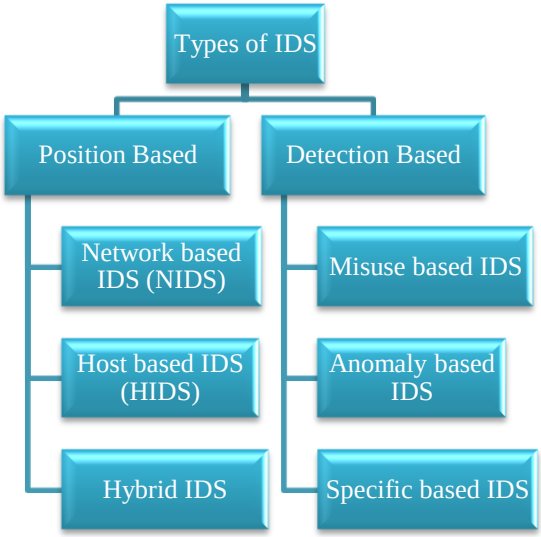

Fig. 3. Types of IDS

D. Dataset and attack signatures used in the wireless attack detection

a. Dataset

There are various investigations that have utilized more established datasets, for example, KDDCUP 99 [23], NSLKDD [18] and numerous specialists demonstrate that these datasets are obsolete now [25], [20]. In this manner, it is fundamental to evaluate new dataset, for instance, Aegean Wi-Fi Intrusion Dataset (AWID) [21]. The real spine of remote IDS is the Aegean Wi-Fi Intrusion Dataset (AWID), a straightforwardly accessible social affair of sets of information in effortlessly passed on form, which contain veritable signs of both run of the mill and intruding 802.11 traffic improvement Restricted to choices like [27] our dataset is orchestrated towards impedance recognizing verification and considerably more particularly obstruction zone in remote structures.

Dataset	Class Labels
AWID-CLS-F-Trn AWID-CLS-F-Tst AWID-CLS-R-Trn AWID-CLS-R-Tst	Flooding, Impersonation, Injection, Normal.
AWID-ATK-F-Trn AWID-ATK-R-Trn	Amok, Arp, Authentication request, Beacon, Cafe latte, Deauthentication, Evil twin, Fragmentation, Probe response, Normal.
AWID-ATK-F-Tst	Amok, Arp, Authentication request, Beacon, Cafe latte, Chop chop, Cts, Deauthentication, Disassociation, Evil twin, Fragmentation, Hirte, Power saving, Probe request, Probe response, Rts, Normal.
AWID-ATK-R-Tst	Amok, Arp, Beacon, Cafe latte, Chop chop, Cts, Deauthentication, Disassociation, Evil twin, Rts Fragmentation, Hirte, Power saving, Probe request, Normal.

Table II: Aegean Wi-Fi Intrusion Dataset (AWID) class distribution

The pursues in AWID do not seem to be only phony but rather segregated ones from a gave WEP ensured 802.11 system. Among the similar ones, the above seems to be the basic wholeheartedly accessible dataset. It is flawed that the remarkable KDD'99 [23] or proportionate sets made for wired conditions won't speedy the formation of upgraded figuring focusing on 802.11 conditions as the two spaces have essentially orchestrated qualities. Everything considered, the AWID dataset may demonstrate a critical contraption for get some information about even on various remote advances such as WiMax [17], LTE [19] and UMTS [24] / settings of elective 802.11 for instance, vehicular systems [22] and work mode [26] as a piece of diverse ambushes depend after taking after models.

The AWID dataset comprises of expansive and diminished datasets. It incorporates separate datasets for preparing (indicated as Trn) and testing (meant as Tst). Each record of the dataset is delegated either ordinary or an explicit interruption type (i.e., class quality of a record is alluded to a sort of interruption or typical system movement). The AWID datasets are principally characterized into two kinds dependent on their class dissemination as abnormal state marked dataset (AWID-CLS) and better grained named dataset (AWID-ATK). The class appropriation of AWID datasets are appeared in Table II [21].

b. Attack Signature [33]

In this section, 15 assaults related with the preparation variant of the lessened AWID dataset (AWID-ATK-RTrn) have been broke down. This undertaking centers in including conceivable snare setups from a hypothetical close by associated see. If irregularities are found in remote framework systems, then the structure perception would be provoked. The same is extended to irregularities also.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Major Attacks of Wi-Fi Network	Sub Attacks of Wi-Fi Network
Flooding	Deauthentication Flooding Attack, Authentication Request Flooding Attack, Beacon Flooding Attack, Probe Flooding Attack
Injection	ARP Injection Attack, Fragmentation Attack
Impersonation	Evil Twin Attack, Cafe Latte Attack

Table III: Types of various Wi-Fi attacks

c. Looding Attacks [33]

Flooding strikes make an astonishing advancement in the organization traces (in their lion's offer) per time unit. The Deauthentication Flooding ambush is assessed as a standout amongst the most grounded DoS strikes in the remote space, yet it is furthermore one of the hardest to conclusively see. In the course, a spurt of Deauthentication diagrams is delivered. Then again, raised periods of such housings may in like way be followed in Amok, Power Saving, Disassociation, Authentication Request Flooding assaults, and in the ARP Injection one when performed inappropriately. Amidst an Authentication Request Flooding strike, Authentication graphs are relied on to demonstrate an epic enlargement Normally, expanded measures of Authentication Requests are available with Amok and Deauthentication Flooding assaults, yet on account of the Authentication Request Flooding the accumulated volume is basically progressively important. A Beacon Flooding strike causes an impossible expansion in the proportion of Beacon follows. The Probe Response Flooding strike impacts in a scene of Probe Response edges. A development of such edges is in like manner seen in the midst of the emulate strikes yet it is generally, significantly milder.

d. Injection Attacks [33]

Injection ambushes frequently root a flood of really encoded data housings of decreased dimension. In ARP type, the assailant is slanted for spreading significant piddling data diagrams which are available for longer period so that proper reaction from the framework system can be drawn out. All through a Fragmentation Attack the interloper blends a movement of short, detached data traces. If beneficial, this strategy consistently does not expend over one moment, regardless if not fruitful an equivalent procedure will be emphasized.

e. Impersonation Attacks [33]

It presents an additional AP for broadcasting a domain based Beacon packets that follows a past significant structure. Each and every impersonation assaults shares a common factor of Beacon frames, that is commonly expanded. Once in a while these ambushes are joined with a surge of non-authenticated plots as a concealed advancement. Caffe Latte ambushes are logically troublesome in nature. In any case, Caffe Latte ambushes will meanwhile saturate encoded Data edges of irrelevant size, much like an ordinary implantation strike, obviously recalling the same from an Evil Twin assault or an ARP Injection is harder.

III. OVERVIEW OF MACHINE LEARNING APPROACHES

In order to find hid bits of information over more than once picking up from data PCs empowers Machine learning. With massive datasets the universe of programming has been changed, which empowers machines to turmoil, re-structure and streamline checks with no other individual. Three classes of machine learning algorithms are depicted in Fig. 4. specifically, classification based learning or supervised learning, clustering based learning or unsupervised learning, & reinforcement learning. Machine learning techniques are consolidated into dual phases, viz., planning & testing. A model is discovered planning data as a reliant one, whereas the readied is related with pass on the craving in the testing stage. This part shows the stray bits of various machine learning techniques as per the clients regard and their latent in dealing with normally troublesome issues [29].

A. Classification based learning

A part of sensible machine learning figuring utilizes a named dataset with classification based learning, somewhere only getting ready model runs with a check. An entire focus of classification based learning is in order to discover the representation of the information attribute space towards the stamp with the target that powerful urge can be made when new information is given. Coordinated learning issues can be likewise asked for into backslide and game plan, where the separation between the two assignments is that the numerical for backslide and names are straight out for request. Portrayal techniques understand how to foresee a class yield for each pushing toward model in context of the arrangement data. By and large standard techniques in this course of action merge Bayesian classifiers [30], k-nearest neighbors (KNN) [46], support vector machine (SVM) [28], decision trees [32] and neural frameworks [34]. Instead of discrete yields, relapse techniques anticipate an industrious regard identifying with every precedent, for instance, assessing the house estimation given its related segment inputs. Standard relapse systems fuse strategic relapse [35], bolster vector relapse (SVR) [36], and Gaussian process for relapse [30].

B. Clustering based learning

The name information helps in teaching directed figuring out of, how to encourage a sensible proportion of advancement that can be used to assess the rightness of the prepared model in various conditions. Be that as it may, an immense amount of marked information is routinely hard to get in readiness. So, unsupervised learning, has been made to locate a

proficient depiction of the data tests with no naming information. A symbolic instance of unsupervised learning is gathering, expressly, to cluster points of reference with the end goal that models in a tantamount get-together have a greater number of likenesses than the models in various social occasions.

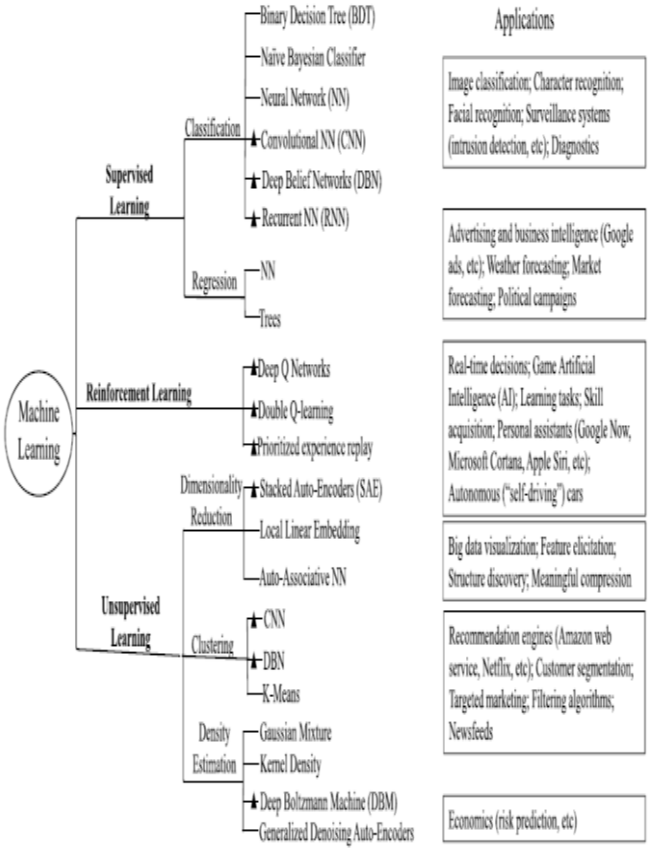

Fig. 4. Categories and Applications of various MachineLearning Approaches

The highlights used for gathering could be either the entire delineation of individual model or the near resemblances among precedents. Standard gathering methodology fuse k-implies [37], progressive bunching [38], range grouping [39], and Dirichlet process [40]. Other than gathering, unsupervised learning also supports dimensionality decrease of high dimensionality space to a lower lossless data. In various circumstances, the crisp information goes with high estimation, which isn't alluring in perspective of a couple of reasons. One reason is the imagined censure of dimensionality [41], that characterizes the testing event experienced when the measurement develops into gigantic. For instance, in order, enhancement and bunching, the model trouble and the amount of basic preparing precedents grow fundamentally with the component measurement. Extra reason is that the contributions of each measurement are normally associated and a few measurements might be demolished with impedance and clamor, which would degenerate the learning execution if not controlled properly. Some standard dimensionality decrease systems fuse direct projection procedures, for example important part examination (PCA) [42], and nonlinear projection techniques, for example, complex learning, neighborhood straight inserting (LLE) [48], and ISOMAP (isometric component mapping) [44].

C. Reinforcement Learning

In reinforcement learning issues, a pro takes in the perfect practices over interfacing with the earth in an experimentation route hoping to increase rewards from nature. Nature is shown as a MDP (Markov choice process) with the entire structure of a Markov system [45].

IV. OVERVIEW OF MACHINE LEARNING DRIVEN IDS IN WIRELESS NETWORKS (WI-FI) OF IOT

Usage of IDS with machine-learning strategy is an occurrence of an individuality dependent IDS [31]. Supervised and unsupervised learning are the two sort of machine learning strategies, that can be used in IDS.

Reference	Application	Problem Statement	Learning Type	Machine Learning Algorithm
Tsang et. al (2006) [47]	Anomaly Intrusion Detection	Feature Extraction & Dimensionality Reduction	Unsupervised	Ant Colony Clustering
Motoda et. al (2002) [43]	Feature Learning in Machine Learning Algorithm	Feature Construction, Feature Extraction & Feature Selection	Unsupervised	Feed Forward Neural Network & Principal Component Analysis
Bostani et .al (2017) [72]	Social Network Analyses	Intrusion Detection	Unsupervised	Optimum Path Forest Graph based algorithm and K-means
Sabhnani et. al (2003) [50]	Intrusion Detection System	Detection of Benign and Malicious Network Traffic Flow	Supervised & Unsupervised	Multiple machine learning classifiers
Kolias et. al, (2011) [52]	Intrusion Detection System	State-of-the-art	Unsupervised	Swarm Optimization Algorithm
Kim et. al (2017) [51]	Hybrid IDS	Anomaly Detection	Supervised & Unsupervised	Fuzzy Inference System & Ant Colony Clustering
Karami et. al (2015) [53]	Content Centric Network	Intrusion Detection	Supervised & Unsupervised	Fuzzy Inference System & PSO algorithm
Kim et. al (2015) [54]	Intrusion Detection System	Detection of Unknown attacks	Unsupervised	Bio-inspired algorithm
Kim et. al (2016) [74]	Intrusion Detection System	Detection of Unknown attacks	Unsupervised	Ant Colony Clustering & Artificial Neural Network

Kim et. al(2014) [56]	Anomaly Intrusion Detection	Botnet Detection	Semi-supervised	Ant Colony Clustering
Udaya et. al (2016) [83]	Intrusion Detection System	Real time wireless network traffic flow attack detection using feature reduction techniques	Supervised & Unsupervised	Information Gain, Chi-Square Statistics, Ada Boost, OneR, J48 Decision tree, Random Tree and Random Forest

Table IV: Summary of IDS Approaches utilizing Machine Learning Techniques in Wireless Networks (Wi-Fi) of IoT [59]

The clustering based learning does exclude a checked data collection for planning, while the classification based learning needs a named data collection. clustering based learning limit is of essential centrality as it empowers a prototype to be endeavored to perceive new strikes without making exorbitant stamps or ward factors.

A blend of two customary methodologies are regularly used to gather an IDS, for example preparing or learning and order as showed in Fig.5. It is hard and costly to get larger piece of marked system affiliation procedures for administered learning in the essential stage. By then element gathering or learning may transform into the outcome regardless. The gathering examination has created as a variation from the norm revelation lately [47]. Bunching IDS, averages a cluster of data examining procedure which sections a course of action of unknown data plots into bunches or gatherings with the end goal that traces inside a gathering resemble each other yet unique to other gatherings' blueprint [47].

Meanwhile, highlight creation builds up the novel highlights to redesign their articulateness, while include expulsion changes the novel highlights into an inventive frame and highlight decision evacuates unnecessary highlights [43]. The order undertaking is a managed procedure to perceive liberal and harmful deals dependent on conveyed data which normally begins from going before stage as exhibited in Fig.5. The pre-taking care of module conventionally involves adjusting and standardization steps. Information standardization is a system to yield all regard extents of each element are proportionate [72]. At that point, the nature of genuine system is having considerate deals impressively greater than threatening deals. This having a place can make it troublesome for the IDS module to take in the key precedents precisely [50]. Subsequently, an adjusting system which makes the dataset with comparable extent for both altruistic and malicious events, is a basic stage for a preparation. On the other hand, novel proportion has been used, for challenging judgments of an IDS in a real time system. Ant Clustering Algorithm (ACA) is a champion unsupervised learning strategy

among the most for the most part used gathering strategies which is begun from swarm knowledge. ACA can have the capacity to find close ideal gathering result without requiring predefined number of gatherings [47]. However, ACA is rarely used in IDS as a tip top strategy for arrangement. Or maybe, ACA is joined with other coordinated figuring, for instance, Support Vector Machine and Self-Organizing Map are allowed to convey enhanced gathering outcome [52]. A unique half and half IDS framework reliant on Ant Clustering Algorithm for preparing stage & Fuzzy Inference System for characterization stage have been proposed Kim et. al, [51]. FIS has been chosen as arrangement stage, in light of the fact that fluffy methodology can diminish the false alarm with higher consistency in choosing interruption happenings [53]. Meanwhile, KKK15 has dissected the equivalent ACA with different classifiers [54] and KHKY16 [74] by using Decision Tree (DT) and Artificial Immune System (AIS) correspondingly.

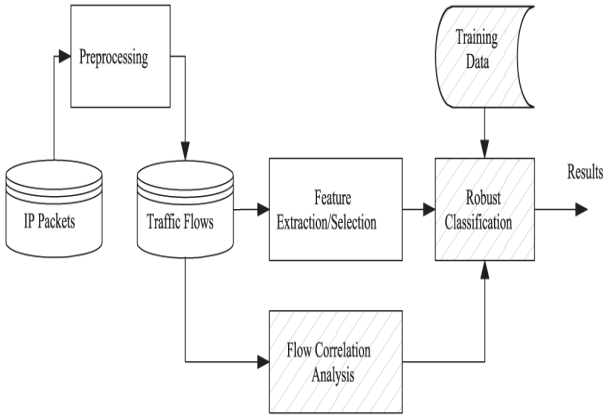

Fig. 5. Typical Scheme of IDS

For the purpose of computational component Artificial Immune System along with Human Inference System have been proposed. Likewise, has researched An improved ACA has been researched by HKY14 by including ATTA-C (Adaptive Time Dependent Transporter Ants Clustering) [56], which is one of the limited methods that have been benchmarked on different datasets, and is before long uninhibitedly accessible under GNU contract [56]. Despite as of late alluded to standard IDSs, distinctive IDS representations pleasing good conditions of Hadoop structure [76] & Software Defined Networking condition has been evaluated [58].

V. OVERVIEW OF DEEP LEARNING APPROACHES

We begin with a short preamble to significant getting the hang of, highlighting the key measures behind figuring strategies in this field, and furthermore key central indicates that lead their thriving. Significant learning is fundamentally a sub-some portion of ML where estimations continuously separate data from rough data, in order to make desires or accept exercises as demonstrated by some goal target. The real advantage of profound learning over customary ML is programmed include extraction, accordingly maintaining a strategic distance from costly hand-created highlight building. We show the connection between profound learning, machine inclining, and man-made consciousness

(AI) at an abnormal state in Fig.6.

Profound taking in at first begins from the developments of Neural Network (NN) figuring [59]. Differing strategies have been related with an unequivocal genuine target to beat the constraints of one subtle layer just in NN. Those strategies utilize predictable unnoticeable layers which are legitimately fell Because of gigantic of frameworks have a place with significant taking in, different significant learning techniques have been assembled subject to their approach [78] as appeared in Table V. Deng [61] recognizes profound learning into three gatherings, unsupervised or generative, managed or discriminative and cross breed. The grouping request relies upon the point of models and strategies, e.g., amalgamation/age or acknowledgment/order.

The requesting of the profound learning techniques is introduced in Fig.7.

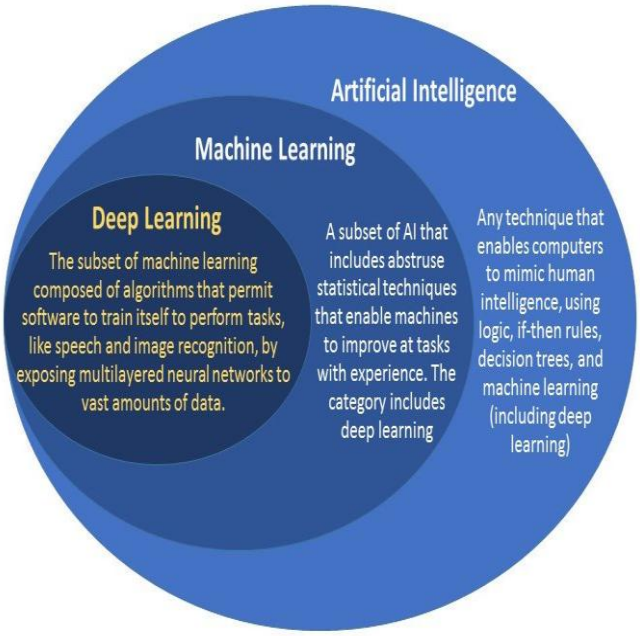

Fig. 6. Venn graph of the connection between deep learning, machine learning and AI

a. Auto Encoder (AE) & Stacked Auto Encoder (SAE)

It is a deep learning approach, with data & yield layers' relative neuron number is a customary Artificial Neural Network (ANN). Then, the center points in the disguised layer are inferring new course of action of features which is low-dimensional. This advancement prompts a limit that can emulate the information after troublesome tallies. AE needs to take in a tinier strategy of information beneficially and can be stacked to make a significant framework. Getting ready outcomes of each covered layer are fell and called as "SAE". This is expected to provide modified features through various profundities [62]. Further, "Denoising Auto Encoder (DAE)" is established for changing strong amended commitment from the tainted one by hullabaloo [63]. DAE might be besides stacked so as to accumulate significant frameworks as well.

b. Boltzmann Machine (BM):

The action of twofold units which are integrated symmetrically is known as Boltzmann Machine with a choices of dynamic structural neural units [64] [65]. Boltzmann Machine without a balanced relationship forms a Restricted Boltzmann machine [64]. By then, one BM result is fell into various BMs, called Deep Boltzmann Machine results by piling various layers of BMs together. If different layers are stacked, layer-by-layer plan, called as Deep Belief Network, which can be utilized for attribute extraction form the unknown data collection [66].

C. Supervised/ Discriminative Learning

Directed or discriminative profound learning is relied upon to perceive a couple of areas of data for configuration arrange [61]. Convolutional Neural Network is an instance of discriminative engineering which uses an interesting structure particularly sensible for picture order. Convolutional Neural Network prepares multiple layer systems with angle plummet to absorb intricate, high-dimensional, nonlinear representation from vast accumulations of information [68]. Neighborhood responsive fields, shared loads, and pooling are the three major needs handled by CNN [69]. One expansive research that viably sent using CNN is AlphaGo by Google [70].

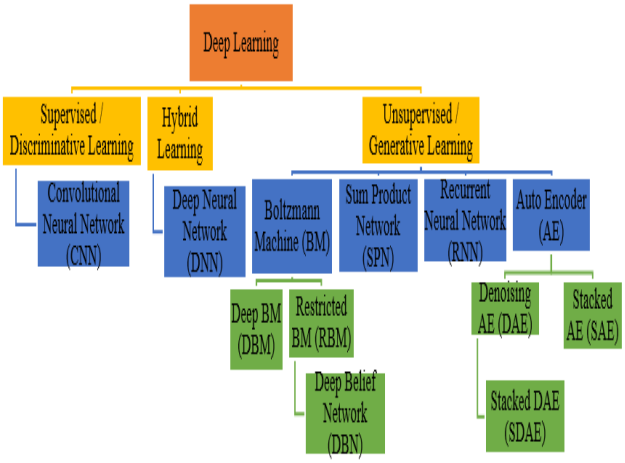

Fig. 7. Classification of Deep Learning Algorithms

Model	Learning Type	Applications	Advantages	Disadvantages
MLP	Supervised, unsupervised, reinforcement	Displaying information with basic connections	Simple arrangement and clear to fabricate	High intricacy, unobtrusive execution and moderate interming ling
RBM	Unsupervised	Removing vigorous portrayal	Can produce virtual examples	Hard to prepare well
AE	Unsupervised	Learning meager and reduced portrayals	Amazing and compelling unsupervised learning	Costly to pretrain with enormous information
CNN	Supervised, unsupervised, reinforcement	Spatial information displaying	Weight sharing; relative invariance	Expensive computation; testing to discover ideal hyper-parameters; demands deep structures in case of complex errands
RNN	Supervised, unsupervised, reinforcement	Successive information displaying	Mastery in catching transient conditions	High model multifaceted nature; angle vanishing and detonating issues

Table V: Summary of Different Deep Learning Techniques [67]

D. Hybrid

Hybrid deep structure joins both discriminative and generative profound learning structures. The half breed design intends to perceive data and moreover discriminative methodology. Deep Neural Network (DNN) is an instance of cross breed design. Regardless, some perplexity terms among DNN and DBN occurs. DBN moreover customs rear spread discriminative learning by way of a "calibrating." So, DBN is a great degree like Deep Neural Network (DNN) [61]. As demonstrated by Deng [65], DNN is described as a multilayer connect with fell totally related concealed layers.

VI. OVERVIEW OF DEEP LEARNING DRIVEN IDS IN WIRELESS NETWORKS (WI-FI) OF IOT

With the expanding prevalence of remote availability, ensuring clients, organize gear and information from pernicious assaults, unapproved access and data spillage winds up vital. Cybersecurity frameworks protect cell phones and clients through firewalls, hostile to infection programming, and an Intrusion Detection System (IDS) [71]. The firewall is an entrance security entryway between two systems. It permits or hinders the uplink and downlink organize traffic dependent on pre-characterized rules. Against infection programming identifies and expel PC infections, worms and Trojans and malwares. IDSs recognize unapproved and noxious exercises or guideline infringement in data frameworks. Each plays out its own capacities to secure system correspondence, focal servers and edge gadgets.

Present day digital security frameworks advantage progressively from profound learning [49], since it can empower the framework to (I) naturally take in marks and examples as a matter of fact and sum up to future interruptions (directed learning); or (ii) recognize designsthat are plainly varied from ordinary conduct (learning without supervision). As a consequence, the pre-characterized rules exertion applicable to separation of interruptions is drastically decreased. Past shielding systems from assaults, profound inclining can likewise assume the job of "assailant", having enormous potential for taking or splitting clients' secret word or data. Outline of these works appear in Table VI. capacities to secure system correspondence, focal servers and edge gadgets.

Present day digital security frameworks advantage progressively from profound learning [49], since it can empower the framework to (I) naturally take in marks and examples as a matter of fact and sum up to future interruptions (directed learning); or (ii) recognize designs that are plainly varied from ordinary conduct (learning without supervision). As a consequence, the pre-characterized rules exertion applicable to separation of interruptions is drastically decreased. Past shielding systems from assaults, profound inclining can likewise assume the job of "assailant", having enormous potential for taking or splitting clients' secret word or data. Outline of these works appear in Table VI.

Irregularity location, which goes for distinguishing system occasions (e.g. assault, sudden access and utilization He utilizes a stacked AE to order organize traffic into 5 types (i.e. authentic, flooding, infusion and pantomime traffic), accomplishing 98.67% generally speaking precision. The AE is likewise abused in [75], where Aminanto and Kim utilize a MLP and stacked AE for highlight choice and extraction, showing surprising execution. Correspondingly, Feng et al. use AE to identify irregular range use in remote correspondence [57]. The investigations propose that the recognition precision can altogether profit by the profundity of AEs.

Appropriated assault identification is likewise an essential issue in versatile system security. Khan et al. concentrate on identifying flooding assaults in remote work systems [77]. They mimic a remote domain with 100 hubs, and misleadingly infuse middle of the road and extreme appropriated flooding assaults to create engineered dataset. Scattered attacks are furthermore analyzed in [60], where the creators focus on an IoT circumstance. Another work in [79] uses MLPs to recognize dispersed forswearing of administration ambushes. The restrictive VAE is proposed for distinguishing the interruption episodes [80]. So as to enhance discovery execution, their VAE gathers missing highlights related with inadequate estimations, which is basic in an IoT domain. The genuine information marks are implanted into the decoder layers to help last characterization. Assessments on the outstanding NSL-KDD [81] show that it accomplishes wonderful exactness in recognizing administration denial, examining, remote-client and client-root assaults, outflanking customary ML strategies by a F1 score 0.18. A novel strategy is proposed in Ref. [82], D-FES, which joins stacked element extraction and weighted component determination systems so as to recognize pantomime assaults in Wi-Fi systemsof information) that don't fit in with a normal conduct, is turning into a key method in IDSs. Numerous specialists put exertion on this territory by misusing the exceptional unsupervised capacity of AEs [73]. For instance, Thing researches highlights of assaults and dangers exist in IEEE 802.11 systems [55].

VII. CONCLUSION AND FUTURE DIRECTIONS

A. Conclusion

In this paper we have investigated an overview of machine learning and deep learning technologies which are being utilized in IDS for the detection of wireless network attacks and system design of effective IDS. Machine learning is a vast and advanced field still relatively immature and definitely not optimized for IDS. The study shows that machine learning based IDS are not efficient enough to handle huge dynamic network traffic flow against Wi-Fi attack detection in wireless networks. On the other hand, machine learning based IDS are constrained to detect the attacks only based on the specific set of features that are pre-determined by some features selection methods. Using this limitation, an intruder can evade the wireless network IDS by crafting the light weighted features that are eliminated by the feature selection methods. To overcome these limitations in the machine learning based IDS, the researchers are shifted to bio-inspired deep learning algorithm called Auto Encoder (AE) to perform an effective feature extraction and dimensionality reduction process in the detection of various wireless network attacks along with the classification and clustering task.

B. Further Research Directions

The problem of wireless attack detection is not yet solved absolutely. Further, the following issues need to be addressed.

1) Most of the existing machine learning and deep learning based IDS systems detect only particular attack type of wireless network, mainly impersonation and also multi-class classification is not performed to classify sub attacks that occur in the wireless networks. Therefore, an IDS not only detecting impersonation attack but also perform detection of flooding and injection attacks along with their sub-attacks has to be implemented by involving multi class classification algorithms in future.

2) Existing deep Stacked Auto Encoder (SAE) based IDS make use of both supervised and unsupervised learning paradigm, which is not feasible to the real time unlabeled wireless network traffic data. Hence, the deep learning based wireless network IDS have to be improved by leveraging unsupervised and semi supervised deep learning algorithms like Boltzmann Machine (BM), RNN, etc., in order to handle huge unlabeled data with dynamic feature learning and feature representation ability.

3) An IDS for distinguishing zero-day assaults but should have capabilities as high and low for discovery rate and false alert rate respectively.

4) Far reaching measure identification as well as counteractive action/moderation procedures are required later on. Then assembling the system with location as well as aversion capacities is normal.

Reference	Application	Problem Statement	Learning Type*	Deep Learning Architecture**
[73]	Applications of cyber security	Malware classification & DoS, probing, R2L & U2R	X & Y	A
[55]	Anomaly detection, attack classification (network- IEEE 802.11)	Flooding, injection and impersonation attacks	X & Y	A
[75]	Wi-Fi impersonation attack detection	Impersonation attack	X & Y	B, C
[57]	Detection of spectrum anomaly	Additive white Gaussian noise	X & Y	C
[77]	Detection of flooding attack (network-wireless mesh)	Flood attack is intermediate but distributed severely	X	B
[60]	Detection of IoT distributed attack	Service Denial, examining, remote- user & user - root	X	B
[79]	Detection of distributed denial - service attack	Distributed denial-both known and unknown	X	B
[80]	Intrusion Detection System for Internet of Things	DoS, probing, R2L & U2R	X & Y	D
[82]	Wi-Fi impersonation detection	Impersonation attack	X & Y	E

*X-Supervised, Y Unsupervised
**A-Stacked AE; B-MLP; C-AE; D-Conditional VAE; E-Weighted Stacked AE

Table VI: Summary of IDS Approaches using Deep Learning Techniques in Wireless Networks of IoT [67

REFERENCES

1. V. N. Index, "The zettabyte era–trends and analysis," Cisco white paper, 2013.
2. Mahmoud Elkhodr, Seyed Shahrestani and Hon Cheung, "Emerging Wireless Technologies in The Internet of Things: A Comparative Study", International Journal of Wireless & Mobile Networks (IJWMN) Vol. 8, No. 5, October 2016.
3. Index, "Global mobile data traffic forecast update, 2010-2015," White Paper, 2011.
4. Sanou, "The World in 2013: ICT facts and figures," International Telecommunications Union, 2013.

5. Vala Afshar, “Cisco: Enterprises Are Leading The Internet of Things Innovation”, August 2017.
6. Christin D, Reinhardt A, Mogre, and Steinmetz, "Wireless Sensor Networks and the Internet of Things: Selected Challenges,” 2009.
7. Xiaoguang, Li, Ke, and Ketai, "The applications of WiFi-based wireless sensor network in internet of things and smart grid," 6th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2011.
8. Jekishan K. Parmar, Ankit Desai, “IoT: Networking Technologies and Research Challenges”, International Journal of Computer Applications (0975 - 8887), Volume 154 - No.7, November 2016.
9. R. Want, B. N. Schilit, and S. Jenson, "Enabling the internet of things," Computer, 2015.
10. Z. Sheng, S. Yang, Y. Yu, A. Vasilakos, J. Mccann, and K. Leung, "A survey on the ietf protocol suite for the internet of things: Standards, challenges, and opportunities," IEEE Wireless Communications, vol. 20, pp. 91-98, 2013.
11. Osseiran, Boccardi, Braun, Kusume, Marsch, Maternia, “Scenarios for 5G mobile and wireless communications: The vision of the metis project,” May, 2014.
12. Kolias, A. Stavrou, J. Voas, I. Bojanova, and R. Kuhn, “Learning internet-of-things security” hands-on”,” IEEE Security Privacy, vol. 14, no. 1, 2016.
13. Farooqi and F. A. Khan, “Intrusion detection systems for wireless sensor networks: A survey,” Communication and networking, 2009.
14. M. Alvarez, N. Bradley, P. Cobb, S. Craig, R. Iffert, L. Kessem, J. Kravitz, D. McMilen, and S. Moore, “IBM X-force threat intelligence index 2017,” IBM Corporation, pp. 1–30, 2017.
15. K. Scarfone and P. Mell, “Guide to intrusion detection and prevention systems (idps),” NIST special publication, vol. 800, no. 2007, 2007.
16. R. Mitchell and I. R. Chen, “Behavior rule specification-based intrusion detection for safety critical medical cyber physical systems,” IEEE Transactions on Dependable and Secure Computing, vol. 12, no. 1, pp. 16–30, Jan 2015.
17. IEEE. 802.16e-2005, I.S. IEEE Standard for Local and Metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems. Nov. 2014. URL: http: / / standards . ieee . org/getieee802/download/ 802.16e-2005.pdf.
18. NSL-KDD. Nsl-kdd data set for network-based intrusion detection systems. http: //www.unb.ca/research/iscx/dataset/iscx-NSL-KDD-dataset.html, 2009.
19. ETSI. Universal Mobile Telecommunications System (UMTS); User Equipment (UE) radio transmission and reception (FDD). Nov. 2014. URL: http://www.3gpp.org/specifications/79-specification-numbering.
20. Maheshkumar Sabhnani and Gursel Serpen. Why machine learning algorithms fail in misuse detection on kdd intrusion detection data set. Intelligent Data Analysis, 8(4): 403–415, 2004.
21. AWID. Awid-wireless security datasets project data set. http://icsdweb.aegean.gr/awid/features.html, 2014.
22. Nikita Lyamin, Alexey Vinel, Magnus Jonsson, and Jonathan Loo. “Real-Time Detection of Denial-of-Service Attacks in IEEE 802.11 p Vehicular Networks”.In: IEEE communications letters 18.1 (2014), pp. 110–113.
23. KDDCUP99. Kdd cup99 data set. http://kdd.ics.uci.edu/databases/ kddcup99/kddcup99.html, 1999.
24. 3GPP. LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (3GPP TS 36.101 version 10.3.0 Release 10). Nov. 2014. URL: http://www.3gpp.org/DynaReport/36-series.htm.
25. M. Tavallaee, E. Bagheri, Wei Lu, and A.A. Ghorbani. A detailed analysis of the kdd cup 99 data set. In Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on, pages 1–6, July 2009. doi: 10.1109/CISDA. 2009.5356528.
26. Rodrigo do Carmo and Matthias Hollick. “DogoIDS: a mobile and active intrusion detection system for IEEE 802.11 s wireless mesh networks”. In: Proceedings of the 2nd ACM workshop on Hot topics on wireless network security and privacy. ACM. 2013, pp. 13–18.
27. Richard Gass, James Scott, and Christophe Diot. “Measurements of in-motion 802.11 networking”. In: Mobile Computing Systems and Applications, 2006. WMCSA’ 06. Proceedings. 7th IEEE Workshop on. IEEE. 2005, pp. 69–74.
28. C. Cortes and V. Vapnik, “Support-vector networks,” Machine learning, vol. 20, no. 3, pp. 273–297, Sep. 1995.
29. Le Liang, and Geoffrey Ye Li, “Towards Intelligent Vehicular Networks: A Machine Learning Framework”, April 2018.
30. G. E. Box and G. C. Tiao, Bayesian inference in statistical analysis. John Wiley & Sons, 2011, vol. 40.
31. C.-H. Tsang and S. Kwong, “Ant colony clustering and feature extraction for anomaly intrusion detection,” Swarm Intelligence in Data Mining, pp. 101–123, 2006.
32. S. R. Safavian and D. Landgrebe, “A survey of decision tree classifier methodology,” IEEE Trans. Syst., Man, Cybern., Syst, vol. 21, no. 3, pp. 660–674, May/Jun. 1991.
33. Constantinos Kolias, Georgios Kambourakis, Angelos Stavrou, and Stefanos Gritzalis, "Intrusion detection in 802.11 networks: empirical evaluation of threats and a public dataset," IEEE Communications Surveys & Tutorials, vol:18.1, pp: 184-208, 2015.
34. LeCun, Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, May 2015.
35. Walker and Duncan, “Estimation of the probability of an event as a function of several independent variables,” Biometrika, vol. 54, no. 1-2, Jun. 1967.
36. Basak, S. Pal, and D. C. Patranabis, “Support vector regression,” Neural Inf. Process. Lett. Rev., vol. 11, no. 10, Oct. 2007.
37. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, “An efficient k-means clustering algorithm: Analysis and implementation,” IEEE Transactions on Pattern Anal. Mach. Intell., vol. 24, no. 7, Jul. 2002.
38. G. Gan, C. Ma, and J. Wu, Data Clustering: Theory, Algorithms, and Applications, 2007.
39. A. Y. Ng, M. I. Jordan, and Y. Weiss, “On spectral clustering: Analysis and an algorithm,” in Proc. NIPS, 2002.
40. Y. W. Teh, “Dirichlet process,” in Encyclopedia of Machine Learning. Springer, 2011.
41. J. H. Friedman, “On bias, variance, 0/1–loss, and the curse-ofdimensionality,” Data Mining Knowl. Disc., vol. 1, no. 1, pp. 55–77, Mar. 1997.
42. I. T. Jolliffe, “Principal component analysis and factor analysis,” in Principal Component Analysis. Springer, 1986, pp. 115–128.
43. H. Motoda and H. Liu, “Feature selection, extraction and construction,” Communication of IICM (Institute of Information and Computing Machinery, Taiwan) Vol, vol. 5, pp. 67–72, 2002.
44. J. B. Tenenbaum, V. De Silva, and J. C. Langford, “A global geometric framework for nonlinear dimensionality reduction,” Science, vol. 290, no. 5500, pp. 2319–2323, Dec. 2000.
45. R. S. Sutton and A. G. Barto, Introduction to reinforcement learning. MIT press Cambridge, 1998, vol. 135.
46. Butun, S. D. Morgera, and R. Sankar, “A survey of intrusion detection systems in wireless sensor networks,” IEEE Communications Surveys & Tutorials, vol. 16, no. 1, 2014.
47. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft, “When is nearest neighbor meaningful?” in Proc. International Conference Database Theory, Jan. 1999.
48. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” Science, 2000.
49. Donghwoon Kwon, Hyunjoo Kim, Jinoh Kim, Sang C Suh, Ikkyun Kim, and Kuinam J Kim. A survey of deep learning-based network anomaly detection. Cluster Computing, 2017.
50. M. Sabhnani and G. Serpen, “Application of machine learning algorithms to KDD intrusion detection dataset within misuse detection context,” Proceedings of the International Conference on Machine Learning: Models, Technologies, and Applications, 2003.
51. Aminanto, H. Kim, K. M. Kim, and K. Kim, “Another fuzzy anomaly detection system based on ant clustering algorithm,” IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 2017.
52. Kolias, G.Kambourakis, and M. Maragoudakis, “Swarm intelligence in intrusion detection: A survey,” Computers & Security, vol. 30, no. 8, 2011.
53. Karami and M. Guerrero-Zapata, “A fuzzy anomaly detection system based on hybrid pso-kmeans algorithm in content-centric networks,” Neurocomputing, 2015.
54. K. M. Kim, H. Kim, and K. Kim, “Design of an intrusion detection system for unknown-attacks based on bio-inspired algorithms,” Computer Security Symposium (CSS), vol. 2015, no. 3, pp. 64–70, 2015.
55. Vrizlynn LL Thing. IEEE 802.11 network anomaly detection and attack classification: A deep learning approach. In IEEE Wireless Communications and Networking Conference (WCNC), pages 1–6, 2017.
56. K. Huseynov, K. Kim, and P. Yoo, “Semi-supervised botnet detection using ant colony clustering,” The 31th Symposium on Cryptography and Information Security (SCIS), 2014.

57. Qingsong Feng, Zheng Dou, Chunmei Li, and Guangzhen Si. Anomaly detection of spectrum in wireless communication via deep autoencoder. In International Conference on Computer Science and its Applications, pages 259–265. Springer, 2016.

58. D. S. Lee and K. Kim, "Improving detection capability of flow-based ids in sdn," KAIST, Department of Computer Science, Thesis Book, 2015.
59. Kwangjo Kim, Muhamad Erza Aminanto, "Deep Learning in Intrusion Detection Perspective: Overview and Further Challenges" IWBIS, 2017.
60. Abebe Abeshu Diro and Naveen Chilamkurti. Distributed attack detection scheme using deep learning approach for Internet of Things. Future Generation Computer Systems, 2017.
61. Deng, "A tutorial survey of architectures, algorithms, and applications for deep learning," APSIPA Transactions on Signal and Information Processing, vol. 3, 2014.
62. Wang, "The applications of deep learning on traffic identification," Blackhat USA, 2015.
63. P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," Journal of Machine Learning Research, vol. 11, no. Dec, 2010.
64. Salakhutdinov and Hinton, "Deep boltzmann machines," Artificial Intelligence and Statistics, 2009.
65. Deng, D. Yu et al., "Deep learning: methods and applications," Foundations and Trends® in Signal Processing, vol. 7, no. 3–4, 2014.
66. Salama, H. Eid, R. Ramadan, A. Darwish, and A. Hassanien, "Hybrid intelligent intrusion detection scheme," Soft computing in industrial applications, 2011.
67. Chaoyun Zhang, Paul Patras, and Hamed Haddadi, "Deep Learning in Mobile and Wireless Networking: A Survey", IEEE Communications Surveys & Tutorials, March, 2018.
68. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, 1998.
69. Nielsen, "Neural networks and deep learning," 2015.
70. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot et al., "Mastering the game of go with deep neural networks and tree search," Nature, 2016.
71. Anna L Buczak and Erhan Guven. A survey of data mining and machine learning methods for cyber security intrusion detection. IEEE Communications Surveys & Tutorials, 18(2):1153–1176, 2016.
72. Bostani and M. Sheikhan, "Modification of supervised OPF-based intrusion detection systems using unsupervised learning and social network concept," Pattern Recognition, vol. 62, 2017.
73. Mahmood Yousefi-Azar, Vijay Varadharajan, Len Hamey, and Uday Tupakula. Autoencoder-based feature learning for cyber security applications, IEEE International Joint Conference on Neural Networks, 2017.
74. Kim, J. Hong, K. Kim, and P. Yoo, "Evaluation of aca-based intrusion detection systems for unknown-attacks, "The 33th Symposium on Cryptography and Information Security, 2016.
75. Muhamad Erza Aminanto and Kwangjo Kim. Detecting impersonation attack in WiFi networks using deep learning approach. In International Workshop on Information Security Applications, Springer, 2016.
76. Huseynov, P. D. Yoo, and K. Kim, "Scalable p2p botnet detection with threshold setting in hadoop framework," Journal of the Korea Institute of Information Security and Cryptology, vol. 25, no. 4, 2015.
77. Muhammad Altaf Khan, Shafiullah Khan, Bilal Shams, and Jaime Lloret. Distributed flood attack detection mechanism