*International Journal of Intelligence Multidisciplinary Engineering Research(IJIMER)* *Vol. 1 (4), Dec 2023, pp: 12-21.*
*ISSN: 2583-9160*

Review Article

# An Empirical Analysis on Intrusion Detection in The Internet of Things

**B. Sowmya[1], Nagendra Muthuluru[2]**
[1]*Research Scholar, Department of Computer Science and Technology, Sri Krishnadevaraya University, Anantapuramu, India.*
[2]*Professor, Department of Computer Science and Technology, Sri Krishnadevaraya University, Anantapuramu, India*
[1]*bsowmya2008@gmail.com,* [2]*nagendramuthuluru@gmail.com*

*Abstract* - **The Internet of Things (IoT) is quickly gaining traction in today's era. However, because of the diversity of hardware capabilities in use, the delicacy of the data stored inside, and related confidentiality concerns, IoT security is becoming a significant research concern and business fields. Many protection mechanisms are inappropriate for Networks owing to their source of energy nature. Thus, it's critical to incorporate 2nd layer defenses. These devices will also need to be tested in several network types and languages to see how they perform successfully. The improvements in IoT intrusion detection methods are the subject of this article. It offers a thorough examination of existing Intrusion Detection Systems (IDS) for IoT technologies, concentrating on architectural types. After that, a suggestion for future IoT-based IDS directions is provided and assessed. Therefore, it demonstrates how standard approaches are ineffective in the IoT sector owing to their inherent flaws. Current IoT intrusion detection research must take a new path to build a safe, reliable, and effective solution for such networking. A scenario is given to demonstrate how security flaws might just be identified passively.**

Keywords: *Intrusion detection systems (IDS), IoT security, wireless sensor networks, universal IDS.*

## 1. INTRODUCTION

The Internet-of-Things (IoT) is a new way of doing things that focuses on creating a pervasive ecosystem of connected devices to improve life style via widespread connection [1]. It is done by the interconnection of sensors and actuators, which allows for intelligent choices based on the findings of a vast amount of data. IoT technologies are projected to provide extraordinary chances for humans to connect [2]. Furthermore, the suggested frameworks for future IoT devices are allowed for all objects to interact directly and share important information [3]. It enables us to create a genuinely designed world where correct data is widely willing to advise the best decision.

The Internet of Things has emerged partly due to the adoption of RFID devices [4]. RFID tags are small, limited radio tags that are used to identify things and animals wirelessly. At the same time, RFID tags permit the wireless intelligent tracking of objects. They are unintelligent and passive, with features that prevent them from logging and comprehending their surroundings [5], collaborating with other devices, and typically stifle the transformation and further analysis of mass of information. Based on the analysis, these interconnecting devices with efficient information analytics, might improve existing universe facilities and services, such gadgets developed from active items to participatory, collaborating, and intelligent devices. Those gadgets integrate sensors with RFID tags to generate wireless devices capable of detecting environment by providing dynamic data while adhering to an initial ideology of low-power and wirelessly connectivity. The options are limited, however, due to the low powered characteristic of these instruments. As a result, the ability to build various large-scale sensors was realized [6] using enabling technologies from wireless computer networks. To cut costs on sensor consumption, it's also necessary to construct these networks as efficiently as possible, which may be done by using an ad-hoc and decentralized communication protocol.

IoT is an emerging vision of upgraded devices and sensors support in the modelling of a novel online digital technologies. It changed our lives as the necessity for globalized access towards heterogeneous device types became it is apparent in every aspect of society. It is renowned for being done through enormous data collecting and interpretation, but increased interconnectedness brings new problems. Computer network safety has long been a key concern. The need to protect sensor-based networks is perhaps higher than ever [7]-[8] since they are employed in many vital infrastructures and applications. The secure data management included inside IoT-based networks is essential to anybody, especially with the emergence of data protection mandating the data acquisition, storage and problems about individual security and above. Furthermore, digital forensics is rapidly becoming an indispensable instrument for law enforcement and anybody seeking to safeguard their

legitimate rights and interests. As a result, a system in order of computer network activities is essential. Security is necessary because the Internet of Things is a new technology that can transform and enhance society.

This article examines a range of intrusion detection systems for the Internet of Things. Through varying configurations of topologies, detection algorithms, and particular threats identified, each solution aims to enhance detection efficacy in various ways and reduce its resource footprint. This research mainly defines the types of architecture specified and the technologies discovered [9]. It is due to the essential feature of IoT associated with the wide range of existing and future technologies that enable it. This study is reviewed to evaluate the existing approaches for IDS adopted for IoT [10]. It results in a concept for a safety system that uses passive network devices to overcome the current security challenges posed by open-medium and restricted equipment. As a result, any number or kind of detection techniques may be added to the system, improving performance and coverage for many applications.

The remainder of the work is designed as section 2 illustrates the advancements in IoT technologies, section 3 reviews IoT threats, and section 4 shows the intrusion detection mechanisms. Section 5 discusses the IDS platform for IoT; section 6 shows the IDS types for detection techniques. Section 7 shows various existing learning approaches for intrusion detection followed by open research challenges in intrusion detection. The summary of the survey is given in section 9.

## 2. IOT TECHNOLOGIES

Despite the rising popularity and implementation of IoT systems, the word IoT only refers to the concept of global connection among intelligent phones. IoT networks are usually made up of disparate, interconnected devices (or "things") and persons in connection. It does not specify how these items should interact. As a result, the Internet of Things is best thought of as a broad phrase encompassing many systems and technologies, including hardware and software, and does not imply any single standard.

Broadband wireless standards are used to drive and build IoT networks (in the majority of cases). RFID is the first technology kinds used in connected systems. WSNs, NFC, 6Lowpan, Zigbee, and other low-power wireless technologies are also utilized, most of which are considered separate wireless mesh technology due to their limited range and bandwidth. Wi-Fi and Bluetooth [11] is used to create a network with somewhat more extensive coverage. Wide-area networks such as 3G, GPRS, 4G, WiMAX, and others may also be used by IoT devices that bridges wired technologies to access the Internet. Other networks are more accessible [12]. While these procedures and devices aren't expressly developed, integration and possibly utilize the wide range of protocols that will need to be considered. The Internet of Things may be conceived as a three-layer model composed of awareness, transfer, and implementation [13]-[14]. The awareness stage includes wearable sensors like RFID and GPS and limited transmission techniques like 802.15.4 and Bluetooth. In contrast, transportation stage has longer-range communications technology like 802.3, IP and 4g. Platforms like cloud solutions for information management and controllers such as transportation systems make up the final deployment step.

Some standards have been built mainly to accommodate low power equipment due to the device's resource restrictions. IEEE 802.15.4, for example, is a reduced body and related to the promotion protocol for resource-constrained wireless devices; 6Lowpan and Zigbee are both based on it [9]. 6lowpan was created as a resource-constrained alternative for networking files that are typically too big for limited resources. 6lowpan is a low body and related to the promotion protocol specifically developed to connect restricted devices connected to the Internet. It compresses IPv6 using IEEE 802.15.4 or other reduced body and media access controls. 6Lowpan is frequently compared to the Routing Algorithm for Low-Power and Lossy Networking (RPL), an inter-reactive routing for limited devices, in the literature. Both are regarded as being the most prevalent IoT-based connectivity setups [15].

This lack of uniformity, in particular, causes problems when seeking to build generic research outcomes to demonstrate precisely what has to be protected. As a result, this work presented an outline of IoT technology, including the network technology utilized and device capabilities. IoT-based networking's are similar to traditional tiered networking stacks, with each layer reliant on the others. Because IoT networks might still be fairly different, it's necessary to think about various IoT protocols. We'll look at vulnerability scanning for protocols designed particularly for IoT networks (like 6lowpan) and quick WSN for in-depth look into IoT enabling interfaces.

## 3. REVIEWS ON IOT THREATS

An assessment of presently available security concerns in the IoT is critically evaluated in this section. These security issues are mostly related to the CIA paradigm. It is essential to guarantee data security because of the extensive data collecting and processing elements of IoT. (Availability, Integrity, Confidentiality). Data assaults may be divided into two categories: passive and aggressive [16]. Some active attacks are focused on the damage or data subversive over the network, whereas attack vectors are preoccupied with data theft or privacy inversion. Because of several intrinsic properties of IoT, security concerns are shared and differ from traditional security problems. Due to the limited nature of these devices, the majority of these issues come from the perceptual layer. According to Barford in [17], all of these safety problems might be regarded as extensions of device energy limits.

Traditional security devices do not suffer from this problem since they are non-mobile and rely on stable (and possibly limitless) power sources. Cryptographic concepts are the baseline approach of information security needs a significant processor and storage for key data storage to be successful [18], much as an available power source can support enormous quantities of data and calculate.

However, difficulties with architecture and application aren't the only thing that makes IoT devices unsafe. Device manufacturers view security as an afterthought, if at all, due to profit-driven businesses and a novel, competitive market [19]. Theft of data is considered the most significant concern due to the devices' predominant sensing nature. Regrettably, the data is commonly dismissed as insignificant. However, this is frequently not the case, as evidenced by the data leakage from smart metres might jeopardize privacy and possibly intrusion detection [19]. A more severe problem is with smart urban, where information privacy violations may result in prejudice, resulting in "an uneven society" [20]. To keep the breadth of this section manageable, it focuses on risks in the IoT Model's perceptual layer. Cyber vulnerabilities to conventional networks are frequently discussed in the literature and primarily involve the transit and app levels.

While modelling IoT-based equipment using network structure (OSI), it's important to remember that so many assaults can start at the application layer, where the holds the view on IoT model is located. These difficulties are comparable [21]. They originate primarily from equipment constraints such as restricted battery life, confined compute mechanisms, and an open mobile wireless environment, all of which make standard security methods hard to execute [22]. Certain schemes have been proposed to address difficulties at this stage, most notably the integration of the security mentioned above mechanisms in a limited form or the addition of physical safety to the machine itself. Due to the restricted nature described above, several of these solutions have been proved to be incorrect. 802.15.4 [22], Bluetooth [23], RFID [24], and Wi-Fi [25] are examples. Furthermore, the above-mentioned techniques do not guard against assaults on the top tiers, which necessitate the use of appropriate IDS [27].

Attributes of the network interfaces used in top collaboration layers like transportation layer, create additional issues: live stream routing or multi-hop, decentralized design, an open system interconnection medium, and others are just a few samples of highly predominant multi-layer insecurities [28]. Regular computer security solutions, in which application-level protocol and applications are frequently protected at the lower ranks by walls or intrusion detection systems (IDS), may provide inspiration to address these challenges. However, traditional computing protection mechanisms consume a lot of resources, and finances on IoT devices are limited to keep device costs down.

As a result, most manufacturers treat security as an afterthought, prioritizing utility above security [29]. Using methods that are further away from teachers' perception is more secure, especially when using features like IPSec for E2E, identity, and integrated encryption based on the extensive available resources on the devices. However, when this traffic flows from less congested to highly congested areas, new solutions are needed. Furthermore, fundamental concerns like DNS spoofing [30], IPv4/v6 based attacks [31], and routing issues [32] continue to plague some systems. Nevertheless, with the deployment of an IDS, they may be more easily discovered than their restricted equivalents.

## 4. INTRUSION DETECTION

This section begins with an outline of IDS with thorough examination of related IDS features to the IoT. Here, IDS are well-known networking security component. However, they are different types of detection rather than prevention, its usage in wireless networking is unrivalled, as preemptive security methods are challenging to deploy [33]. IDS comes in two flavours: venue and web. Host-based systems keep track of systems (API calls, disc activity, memory consumption, and so on), whereas p2p systems keep track of internet activity and messages. Generally, IDS looks for indicators of attack in conduct (network traffic/ host activity), operating under the notion that legitimate and malicious behaviour is different [34].

An IDS' effectiveness may be measured using two different measures. False positives and false negatives are terms used to describe these situations. When regular traffic is labelled as unlawful, a false positive happens, and when illegal activities are not identified, false-negative results. Although, data sets are scarce, usefulness of assessing success is debatable [35]. The authors have proposed many different methods for creating various types of IDS. Because of both logo database and anomaly modelling, the bulk of these requires a lot of resources [36]. Furthermore, to maintain the datasets or model correct, each detection mentioned above requires non - periodic updates. Both of the detection above approaches is not well appropriated to the restricted resources because of this naturally high resource [37]. Various assault detection strategies are discussed widely. The review categorizes the work based on the sort of architecture used, emphasizing the technology discovered. Misuse, abnormal, restriction, or combination is the most common detection kinds [38].

A database of common threats is used in misuse investigative techniques. This database compares activities like traffic on a network and system-level operations to signatures. If a match is found, the activity is marked as suspicious. Continually checking for known vulnerabilities or discovering shellcode in ethernet frames are examples of abnormal network behaviour. Misuse detection is quite good at identifying known attacks (few false positives) but not so

well at detecting new attacks (high false negatives). It is because further assaults have no mark. Furthermore, on limited devices, keeping and updating the signature database is impracticable.

Anomaly detection approaches use a different system, building a model of normal behaviour that can then be matched to actual activity, with any disparities being marked as suspect. For instance, the model may track time and utilization of system programmes; if an application is utilized from typical hours, abnormal behaviour is highlighted. With wireless LAN activity models, on the other hand, when the server is detected connecting to an email or service that is not normal, harmful behaviour is reported again. Outlier detection approaches excel in detecting new assaults in situations when abuse detection methods would generally fail, resulting in a low true alarm rate. They do, however, have a high incidence of false positives when these models are not updated regularly. False positives may occur due to the changing nature of mobile technology. Furthermore, upgrading the models regularly might be resource expensive, burdening devices with limited resources [39].

Anomaly and malicious behaviour are combined in specification-based methods. As previously, this entails using a pre-defined model to detect abnormal behaviour. Because of the human contact, this approach is favourable in terms of enhanced accuracy. Still, it creates a delay in the formation of a signature, causing the operation to be delayed. On the other hand, the action must be certified as malevolent by a human participant [40].

Any combination of those above is used in hybrid detection methods. When there is a flaw in one method's efficacy is offset by the positives [41]. As said before, the Internet of Things (IoT) encompasses a wide range of devices. For a variety of reasons, classifying work according to technology type might be challenging. They are frequently owing to the solution's ambiguity, such as a non - compliance and a purely theoretical suggestion. Many works list WSNs, which may be made up of many protocols. In contrast, others mention a single device type like mobile (laptops, smartphones) standard/ multi-layers or atomic specifications, such as Wireless.

Table 1 Comparison of various attack types and their consequences

| Attack | IoT features | Attack outcomes | Type | Samples |
|---|---|---|---|---|
| Device jamming | Open wireless intermediate, embedded plan, | DoS | Active | Reactive, random, deceptive, constant |
| Network sniffing | Open wireless intermediate, insecure direction-finding, decentralization. | Data disclosure, Privacy incursion | Passive | --- |
| Battery exhaustion | Embedded design, open wireless intermediate | Denial of service, data exposé | Active | Traffic flooding |
| Device cloning | Exterior Deployment, Embedded Design | Data revelation, advanced cryptographic attack | Active/Passive | --- |
| Side-channel analysis | Multi-hop networking, decentralizations | Denial of service, data misdirection, data disloyalty | Passive | --- |
| Routing attacks | Open wireless intermediate, insecure routing, decentralization | Denial of service, data misdirection, data sedition | Active | discriminatory forwarding, slight package alteration, sinkhole |
| Cryptographic attacks | Open wireless intermediate constrained income | Secured data disclosure, | Active/Passive | Brute force |

## 5. IDS TYPES BY MONITORED PLATFORM

### 5.1 Network-based IDS (NIDS)

A NIDS is used to identify and defend all nodes against attacks in the network through connectivity. This type of IDS analyses and models traffic to detect routine business and possible suspects, as intrusions generally occur in irregular patterns. They are made up of a series of sensors installed at different networking points to observe traffic. Every sensor does regional analysis and alerts a central management console to any questionable behaviour. A NIDS may collect and analyze full sent frames, including IP addresses, payloads and ports. NIDS are useful to monitor IP traffic. As long as the IDSs are correctly positioned, an

extensive network is observed with only a few deployed IDSs. This IDS is typically easy to install on a network and is deemed reliable against faults [42]. However, they have certain drawbacks, such as the difficulties digesting all packets from an extensive, overburdened network. As a result, they may miss an assault conducted during moments of high traffic. Furthermore, several benefits of NIDS do not apply to contemporary networks relies on switching, which divide the network and require monitor ports to function effectively. In a switch port, port mirroring or spanning provides a comprehensive picture, resulting in overhead. Furthermore, several benefits of NIDS do not apply to contemporary network model-based on switching, which divide the network and require monitor interfaces to function effectively. In a switch port, port mirroring or spanning provides a comprehensive picture, resulting in expense.

5.2 Host-based IDS (HIDS)

An IDS that operates on specific hosts is known as a HIDS. Its primary goal is to observe host events and identify suspicious actions, such as attacks of the observed computer or attacks against the network host on which it runs. Because this sort of IDS is meant to work with single host, it can perform things that a NIDS can't, such as combining code review, identifying memory leaks, monitoring calls, access abuse, power abuse, system log analysis, and so on. Because they need the software installed on the hosts, these systems can be categorized as agent-based [43]. This IDS assesses the network's security by looking at Linux kernel log files, access logs, and server logs, for example. Because they are placed at the destinations, they protect against attacks that NIDS need not identify, i.e. those relying on encryption protocols. Another advantage of HIDS over NIDS is that an attack is actually success or failure may be assessed quickly. Fig 1 and Fig 2 depicts the representation of NIDS and HIDS.

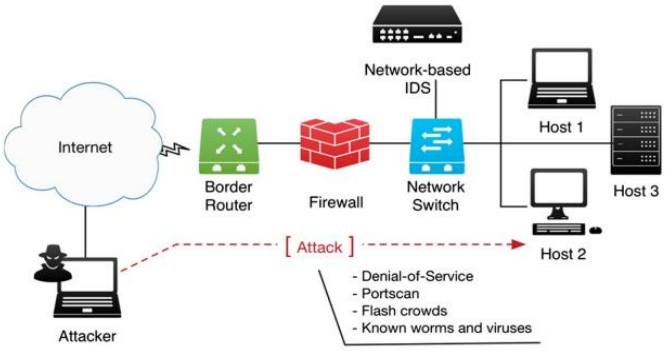


Fig 1 Network-based IDS

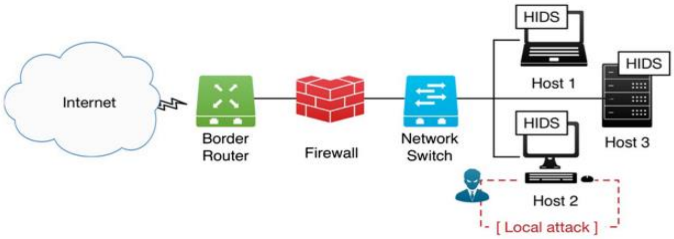


Fig 2 Host-based IDS

5.3 Hybrid IDS

Hybrid IDSs are created by integrating the functions of both NIDS and HIDSs and taking into account data supplied by host events and network segments [44]. These solutions combine the advantages of both techniques while removing the many disadvantages. On the other hand, hybrid systems aren't necessarily superior because different IDS systems analyze data and seek invasive behaviour in distinct manners, integrating and coexisting multiple technologies in a unified system safely and efficiently.

## 6. IDS TYPES FOR DETECTION TECHNIQUE

This section discusses four different detection approaches like signature-based, anomaly-based, and specification-based and hybrid approaches.

6.1 Signature-based (misuse detection)

It is also known as knowledge-based or abuse detection, assesses network activity with catalog of known signatures. The IDS raises the alarm if an attempted match a signature. This procedure guarantees efficient monitoring with few false alarms and high accuracy in identifying and classifying anomalies, making it more straightforward for network managers to take preventative or remedial action. Unknown abnormalities, or minor changes in known assaults, cannot be identified since any activity not recognized by the IDS knowledgebase is deemed ordinary [45]. As a result, signature-based IDSs need to keep their knowledge databases updated regularly. To guarantee that all possible versions of an attack are covered, signatures are specified. They also don't match non-malicious behaviours, which might be difficult.

6.2 Anomaly-based detection

Oddity approaches, also termed as anomaly or username detection, create a baseline profiling representing normal/expected network behaviour. Any detected divergence from this profile is deemed abnormal. The majority of the data used to generate this profile comes from statistics and historical network traffic data. When the user considers Internet for certain time during work hours, this sort of detection is a typical case. Assume this user is management at a firm under the scrutiny of an anomaly-based IDS. These IDS establishes regular profile, and it

began using it as necessary for either because using the Internet on the last day of that week. While the detection is running, the boss needs to access the Internet to post a last-minute report, which is strange behaviour [46]. IDS' unusual response to this odd behaviour is to deny that user Internet service, which would be reasonable if this were not an exception; nevertheless, this would be considered false positives. The most widely utilized IDS detection approach is anomaly detection. It is owing to their capacity to identify both known and undiscovered assaults and aberrations, as identification is dependent on finding odd patterns, making this approach more reactive than handwriting techniques. It also aids in discovering new types of abuse and behaviour and the creation of new signatures for abuse detection methods.

6.3 Specification-based

Anomaly detection methods identify the consequence of aberrant behaviour, whereas misuse detection systems recognize already known abnormal conduct, as stated in [63,66]. As a result, specification-based approaches were developed to make use of the advantages of both techniques. As a result, these IDSs create requirements and restrictions by hand to describe regular network activity. This approach entails determining a programs or protocol's proper functioning and tracking its performance using defined conditions. This detection method is not as commonly used as the others listed in this article, owing to its higher design complexity and the fact that it is limited in its planned use as it is targeted, for instance, at being single application. Fig 3 depicts the state of attack.

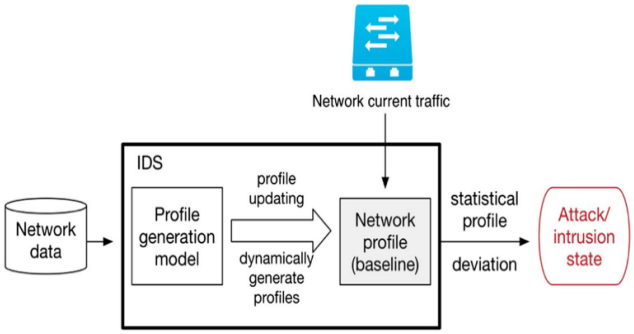


Fig 3 Attack State

6.4 Hybrid techniques

Compound detection, also known as hybrid IDSs, uses a combination of misuse, anomaly, and specification detection approaches. Vijayanand in [47] developed a hybrid IDS in which they generated a networking profile termed digital signature utilizing flow analysis (DSNSF) to identify unknown abnormalities in network traffic. The detected unusual activity was then categorized as DDoS, DoS, flash crowd or port scan assault using pre-loaded fingerprints. These systems might, for example, it relies in regular attack activity and network profile. The author used a combination of reference implementation and anomaly-based approaches to overcome the constraints of the former. A strategy for the autonomous creation of normal and abnormal behavioural requirements as varying patterns identified using anomaly-based machine learning techniques eliminates user knowledge. Table 2 depicts the network anomaly detection.

Table 2. Network Anomaly Detection

| Classification Approach | Types | Merits | Demerits |
|---|---|---|---|
| Data source / Monitored platform | NIDS | Both inward and outgoing network traffic is monitored. | Receiving all items from an extensive and overburdened network is difficult. |
| | | Identify network-specific attacks like DoS attacks. | Failure to identify assaults undertaken during high-traffic periods |
| | | Identify worms and viruses, port scan and flash crowds | Analysis of encrypted packets is not possible. |
| | Host-based (HIDS) | Detect suspicious local activities | In today's massive network, the demand for more sensors is expensive. |
| | | Because they are placed on the endpoint, known attacks are based on encrypted data. | Incomplete network depiction |
| | Hybrid | Privilege abuse, cushion overflows. | Because they are operative, they require support for a variety of operating systems. |
| | | Cumulative benefits of both approaches | Obtain many ways to collaborating and coexisting in an identical system. |
| Detection technique | Misuse detection | Prevail over many drawbacks High discovery accuracy | Unable to spot unidentified abnormalities. Building and updating signatures is a challenging and time-consuming job. |

| | | | |
|---|---|---|---|
| | Adopts prior-knowledge attack database | Low false alarm tempo | High false positives and negatives |
| | Anomaly detection | Become aware of both known and indefinite anomaly | Less efficient in the lively net environment |
| | Side view representing regular network behaviour | Realize new attacks (and use on signature-based IDSs) | Demand time and property to put up the profile |
| | Specification-based | No require for prior knowledge | Complexity |
| | Deposit of limitations to demonstrate and scrutinize the business of protocol | Unknown attacks find | Elaborates terms and constraints<br>It is time-consuming and costly. |
| | Hybrid | Low false positive rates.<br>Anti to understated attack.<br>Combine repayment of approaches.<br>Surmount various disadvantages | Open to the proper operation of a program or protocol.<br>Acquire distinct approach to coexist and interoperate in single system. |

## 7. EXISTING LEARNING APPROACHES

Peiying et al. [48] suggested a Nave Bayesian method for real-time detection of black holes, preferential relaying, and DDoS assaults. As a result of the system's monitoring of packets delivered by nodes, their behaviour is examined to detect anomalies. The properties are similar to that customarily distributed and then uses a standard usual likelihood method to compute the chance of a sample being in a class. Belal et al. [49] utilized a Nave Bayesian method. Still, it is paired with time slicing function to leverage the connection among time and packet headers because network traffic fluctuates at different times, and traffic does not occur. Support Vector Machine (SVM) is another categorization approach that is also utilized in analytical thinking. SVM is a supervised training model provided with a raster images (RBF). The lack of local extremum, incompleteness, and capacity check are attained by operating on the border. Classification methods with high generalization can accurately identify the class of fresh input from same region during the occurrence of learning process.

To address unbalanced class dispersion scenarios and decrease the prevalence of assaults in the traffic data to train an SVM, Keshgary et al. [50] presented a unique technique to giving autonomous labelling to regular traffic. Robust SVM model is used to modify the uncontrolled one-class SVM. Their objective was to reduce the sensitivity of the judgment border to data outliers. The author has developed an efficient IDS based on enhanced features in an SVM. The SVM is used with the exponential maximal densities ratios transformation (LMDRT) which is a feature transmission approach in their framework. The Classifier is improved by using the fresh and succinct dataset to train it. The authors obtained a rapid training time, excellent accuracy and heuristic detection, and minimal false detection contingents by assessing the framework using the widely known NSL-KDD dataset.

Anwar et al. [51] presented an IDS built on the regression superports support vector, a variation of the conventional SVM classifier (LS-SVM). Compared to the standard SVM, this change is more susceptible to outliers and turbulence in the testing set. There are two steps to their decision-making process. The first stage can decrease the dataset dimension by applying an optimal allocation strategy to choose examples based on data variability. The LS-SVM is then fed these sample sizes in the following stage. The method was designed to function with both static and incremental data, and it yielded good outcomes. Prabha Devi et al. [52] used an ANN model to create an intelligent agent that can determine whether the fundamental pattern of data sets are normal or aberrant and categorize them into new and unseen records. Feedforward back propagation (BP) algorithms are used to achieve this aim. They're in charge of providing the neural network with vectorized inputs, comparing the estimated and predicted output and then adjusting the weighted ANN nodes to approach the result. This technique has proven to be robust in throughput and minimal in computational overhead after a few tests. Crossler et al. [53] created an ensemble distributed classifier for NIDS based on the latest tree-level technique for aggregating the multiple learners' choices. The method is based on neural net groups created using genetic programming. It generates a program using dynamic method to show how to integrate projections of the constituent networks to get solid ensembles forecast. However, there are various disadvantages:

- Excessive use of resources
- Inability to detect unknown abnormalities in the absence of suitable training data.
- The employment of neural networks may result in over-fitting.

- For large datasets, the choice of the set called is sluggish.
- True performance might be challenging to obtain in some instances.

## 8. OPEN ISSUES

Within the subject of outlier detection, there are several problems [54] – [55]. This part tries to summarize the most critical open topics discovered during the writing of this article and consider those that have received the most excellent attention from scholars.

Rationality as a notion: It is among the essential phases in developing a network anomaly detection system. The topic of "how to construct a precise concept of normality?" has prompted many scholars to set various answers over time. It is the most significant problem in anomaly detection, and it has yet to be fully solved. It is the objective of many of the works covered in this survey.

Adaptability: Anomalies change when new anomalies are developed, or existing ones are enhanced to bypass present detection techniques. As a result, to react to such developments, IDSs must be updated regularly, which is not a simple process.

Update your profile in real-time: The profile data must be updated with fresh data when an unknown threat is identified and handled by NIDS. However, constantly performing such changes without sacrificing speed or causing conflicts is a problem.

Data with a lot of noise: Normal changes in datasets might be misinterpreted as anomalies if they are not adequately characterized, which is an issue when constructing a profile. Furthermore, neither public nor private databases always make information obvious.

Rates of false alarms: Another issue is reducing false alarms as much as possible. However, it is still impossible to eliminate them and construct a 100% trustworthy IDS. That is still a difficult task.

Standard datasets: Only a few publicly available incursion databases with sufficient information regarding assaults are accessible; moreover, none of them is common assessment dataset for outlier detection. The absence of dependable public datasets can accurately mimic network settings remains an issue.

Monitoring in real-time: As Online data increases annually, the quantity of traffic created by computer networks is continually rising. As a result, developing an effective network management method in real-time has proven challenging.

Complexity: As researchers attempt to address all of the abovementioned issues, the system's intricacy grows as new techniques and approaches are added and mixed. Furthermore, the complexity of network infrastructures adds to the continuation of challenge in terms of data collection and preparation.

## 9. CONCLUSION

This literature study targeted to give theoretical knowledge of intrusion detection problem and its elements. It also sought to provide a comparison of the many approaches explored to solve this issue. There was a debate on what an aberration is and how to recognize its most typical appearances. Anomalies are classified as point abnormalities, collect anomalies, or context oddities, depending on their nature. However, they are separated into operational risks, flash crowds, measurement abnormalities, and attacks according to their causative component. Their accurate detection plays a vital role in creating an IDS that can concentrate on the significant restrictions associated with each type of occurrence. The definition of an IDS and its many kinds are also explored. The best IDS type to create relies on whether for local or wide-area wireless traffic or identifying unknown abnormalities while sacrificing precision. Because it is dynamic and detects both known and unknown abnormalities, anomaly-based identification is the most popular IDS. This research looked at several publications to get a comprehensive picture of what has been done in anomaly detection and what may be done better. There are several ways to solve the anomaly detection problem, ranging from simple techniques to sophisticated systems. However, each methodology has its own set of benefits and downsides.

Furthermore, the results of this study highlighted the most important unresolved challenges in the area. The lack of a standard and updated labelled dataset has been identified as a significant gap. Therefore, creating a public database that covers numerous anomalies and actual traffic patterns of various network architectures. To summarize, there are still some unique challenges to enhance intrusion detection systems' efficacy and practicality. Still, there are several potential suggestions for academics to follow in any further study on the subject.

## REFERENCES

[1]. C. Wu et al., "A Hybrid Intrusion Detection System for IoT Applications with Constrained Resources," International Journal of Digital Crime and Forensics (IJDCF), vol. 12, no. 1, pp. 109-130, 2020.

[2]. Soe, Y. Feng, P. I. Santosa, R. Hartanto, and K. Sakurai, "Towards a Lightweight Detection System for Cyber Attacks in the IoT Environment Using Corresponding Features," Electronics, vol. 9, no. 1, p. 144, 2020.

[3]. You, K. Yim, V. Sharma, G. Choudhary, R. Chen, and J.-H. Cho, "On IoT Misbehavior Detection in Cyber-Physical Systems," in 2018 IEEE 23rd Pacific Rim International Symposium on Dependable Computing (PRDC), 2018, pp. 189-190: IEEE.

[4]. Amouri, V. T. Alaparthy, and S. D. Morgera, "A Machine Learning-Based Intrusion Detection System for Mobile Internet of Things," Sensors, vol. 20, no. 2, p. 461, 2020.

[5]. Qadri, R. Ali, A. Musaddiq, F. Al-Tudjman, D. W. Kim, and S. W. Kim, "The limitations in the state-of-the-art counter-measures against the security threats in H-IoT," Cluster Computing, pp. 1-19, 2020.

[6]. Al-Hamadi and I. R. Chen, "Adaptive network defence management for countering a smart attack and selective capture in wireless sensor networks," IEEE Transactions on Network and Service Management, vol. 12, no. 3, pp. 451-466, 2015.

[7]. Ng, M. B. I. Reaz, and M. A. M. Ali, "A review on the applications of Petri nets in modelling, analysis, and control of urban traffic," IEEE Transactions on Intelligent Transportation Systems, vol. 14, no. 2, pp. 858-870, 2013.

[8]. Andrade and B. Nogueira, "Dependability evaluation of a disaster recovery solution for IoT infrastructures," The Journal of Supercomputing, pp. 1-22, 2018.

[9]. Hashim, F., Munasinghe, K. S., & Jamalipour, A. (2010). Biologically inspired anomaly detection and security control frameworks for complex heterogeneous networks. IEEE Transactions on Network and Service Management, 7, 268–28.

[10]. Balakrishnan, S. M., & Sangaiah, A. K. (2017). MIFIM—Middleware solution for the service-centric anomaly in future internet models. Future Generation Computer Systems, 74, 349–365.

[11]. Lu, S., Wang, X., & Mao, L. (2014). Network security situation awareness based on network simulation. In 2014 IEEE workshop on electronics, computer and applications (pp. 512–517)

[12]. Hosseini Bamakan, S. M., Wang, H., & Shi, Y. (2017). Ramp loss K-support vector classification-regression: A robust and sparse multi-class approach to the intrusion detection problem. Knowledge-Based Systems, 126, 113–126.

[13]. Lakhina, A., Crovella, M., & Diot, C. (2004). Diagnosing network-wide traffic anomalies. In ACM SIGCOMM computer communication review (Vol. 34, p. 219)

[14]. Patcha, A., & Park, J.-M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. Computer Networks, 51, 3448–3470

[15]. Marnerides, A. K., Schaeffer-Filho, A., & Mauthe, A. (2014). Traffic anomaly diagnosis in Internet backbone networks: A survey. Computer Networks, 73, 224–243

[16]. Ahmed, M., Naser Mahmood, A., & Hu, J. (2016). A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, 19–31.

[17]. Barford, P., & Plonka, D. (2001). Characteristics of network traffic flow anomalies. Proceedings of the 1st ACM SIGCOMM workshop on internet measurement (pp. 69–73)

[18]. Mouton, F.,Malan,M.M., & Venter, H. S. (2013). Social engineering from a normative ethics perspective. In Information security for South Africa, 2013 (pp. 1–8).

[19]. Peng, T., Leckie, C., & Ramamohanarao, K. (2007). Survey of network-based defence mechanisms countering the DoS and DDoS problems. ACM Computing Surveys, 39, 3

[20]. Raza, M., & Haider, W. (2012). A survey of password attacks and comparative analysis on methods for secure authentication. World Applied Sciences Journal, 19, 439–444.

[21]. Jadidi, Z., Muthukkumarasamy, V., Sithirasenan, E., & Singh, K. (2015). Flow-based anomaly detection in big data. In Network big data (pp. 257–279).

[22]. Zhang, Y., Fang, B., & Luo, H. (2010). Identifying high-rate flows based on sequential sampling. IEICE Transactions on Information and Systems, E93–D, 1162–1174

[23]. Butun, I., Morgera, S. D., & Sankar, R. (2014). A survey of intrusion detection systems in wireless sensor networks. IEEE Communications Surveys & Tutorials, 16, 266–282

[24]. Meng Hui, L., & Jones, A. (2008). Network anomaly detection system: The state of the art of network behaviour analysis. In International conference on convergence and hybrid information technology 2008. ICHIT '08 (pp. 459–465).

[25]. Stakhanova, N., Basu, S., & Wong, J. (2010). On the symbiosis of specification-based and anomaly-based detection. Computers & Security, 29, 253–268.

[26]. Lakhina, A., Crovella, M., & Diot, C. (2005). Mining anomalies using traffic feature distributions. ACM SIGCOMM Computer Communication Review, 35, 217

[27]. Hamdi, M., & Boudriga, N. (2007). Detecting Denial-of-Service attacks using the wavelet transform. Computer Communications, 30, 3203–3213.

[28]. Pascoal, C., Rosario de Oliveira, M., Valadas, R., Filzmoser, P., Salvador, P., & Pacheco, A. (2012). Robust feature selection and robust PCA for internet traffic anomaly detection. In INFOCOM, 2012 Proceedings of IEEE (pp. 1755–1763).

[29]. Fernandes, G., Carvalho, L. F., Rodrigues, J. J. P. C., & Proença, M. L. (2016). Network anomaly detection using IP flows with principal component analysis and ant colony optimization. Journal of Network and Computer Applications, 64, 1–11

[30]. Fernandes, G., Rodrigues, J. J. P. C., & Proença, M. L. (2015). Autonomous profile-based anomaly detection system using principal component analysis and flow analysis. Applied Soft Computing, 34, 513–525

[31]. Huang, T., Sethu, H., & Kandasamy, N. (2016). A new approach to dimensionality reduction for anomaly detection in data traffic. IEEE Transactions on Network and Service Management, 13, 651–665

[32]. Bang, J., Cho, Y.-J., & Kang, K. (2017). Anomaly detection of network-initiated LTE signalling traffic in wireless sensor and actuator networks based on a Hidden semi-Markov Model. Computers & Security, 65, 108–120.

[33]. Rajasegarar, S., Leckie, C., & Palaniswami, M. (2014). Hyperspherical cluster-based distributed anomaly detection in wireless sensor networks. Journal of Parallel and Distributed Computing, 74, 1833–1847.

[34]. Carvalho, L. F., Barbon, S., Mendes, L. S., & Proença, M. L. (2016). Unsupervised learning clustering and self-organized agents applied to help network management. Expert Systems with Applications, 54, 29–47.

[35]. Su, M.-Y. (2010). Discovery and prevention of attack episodes by frequent episodes mining and finite state machines. Journal of Network and Computer Applications, 33, 156–167.

[36]. Catania, C. A., Bromberg, F., & Garino, C. G. (2012). An autonomous labelling approach to support vector machines algorithms for network traffic anomaly detection. Expert Systems with Applications, 39, 1822–1829

[37]. M. A. Hayes and M. A. M. Capretz, "Contextual Anomaly Detection in Big Sensor Data," in 2014 IEEE International Congress on Big Data, Jun. 2014, pp. 64–71.

[38]. Chalapathi and S. Chawla, "Deep Learning for Anomaly Detection: A Survey," arXiv:1901.03407 [cs, stat], Jan. 2019, arXiv: 1901.03407. [Online]. Available: http://arxiv.org/abs/1901.03407

[39]. Bose, B. Kar, M. Roy, P. K. Gopalakrishnan, and A. Basu, "Adepos: anomaly detection based power saving for predictive maintenance using edge computing," in Proceedings of the 24th Asia and South Pacific Design Automation Conference. ACM, 2019, pp. 597–602.

[40]. Abraham and A. Chuang, "Outlier Detection and Time Series Modeling," Technometrics, vol. 31, no. 2, pp. 241–248, May 1989. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/00401706.1989.10488517

[41]. Y. Zhang, N. Meratnia, and P. J. Havinga, "Outlier detection techniques for wireless sensor networks: A survey." IEEE Communications Surveys and Tutorials, vol. 12, no. 2, pp. 159–170, 2010.

[42]. Zhu and S. Sastry, "Revisit Dynamic ARIMA Based Anomaly Detection," in 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, Oct. 2011, pp. 1263–1268

[43]. Sana Ullah Jan, Van-Hiep Vu, and Insoo Koo. Throughput maximization using an SVM for multi-class hypothesis-based spectrum sensing in cognitive radio. Applied Sciences, 8(3):421, 2018

[44]. Sana Ullah Jan, Young-Doo Lee, Jungpil Shin, and Insoo Koo. Sensor fault classification based on support vector machine and statistical time-domain features. IEEE Access, 5:8682–8690, 2017.

[45]. Sung-Min Lee, Bok-deuk Jeong, and Sang-bum Suh. Method of intrusion detection in terminal device and intrusion detecting apparatus, April 15 2014. US Patent 8,701,188.

[46]. Doohwan Oh, Deokho Kim, and Won Woo Ro. A malicious pattern detection engine for embedded security systems in the Internet of things. Sensors, 14(12):24188–24211, 2014.

[47]. Vijayanand, D Devaraj, and B Kannapiran. A novel intrusion detection system for wireless mesh network with hybrid feature selection technique based on ga and mi. Journal of Intelligent & Fuzzy Systems, 34(3):1243– 1250, 2018

[48]. Peiying Tao, Zhe Sun, and Zhixin Sun. An improved intrusion detection algorithm based on ga and SVM. IEEE Access, 6:13624–13631, 2018.

[49]. Belal Sudqi Khater, Ainuddin Abdul Wahab, Mohd Idris, Mohammed Abdulla Hussain, Ashraf Ahmed Ibrahim. A lightweight perceptron based intrusion detection system for fog computing. Applied Sciences, 9(1):178, 2019

[50]. M Keshtgary, N Rikhtegar, et al. Intrusion detection based on a novel hybrid learning approach. Journal of AI and Data Mining, 6(1):157–162, 2018

[51]. Anwar, Z. Inayat, M. F. Zolkipli, J. M. Zain, A. Gani, N. B. Anuar, M. K. Khan, and V. Chang, "Cross-VM cache-based side-channel attacks and proposed prevention mechanisms: A survey," Journal of Network and Computer Applications, vol. 93, pp. 259–279, Sep. 2017.

[52]. Prabhadevi and N. Jeyanthi, "Distributed Denial of service attacks and its effects on Cloud environment- a survey," in The 2014 International Symposium on Networks, Computers and Communications, Jun. 2014, pp. 1–5.

[53]. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, "Security, privacy and trust in Internet of Things: The road ahead," Computer Networks, vol. 76, pp. 146–164, Jan. 2015.

[54]. Sikder, H. Aksu, and A. S. Uluagac, "6thsense: A Context-aware Sensor-based Attack Detector for Smart Devices," 26th USENIX Security Symposium (USENIX Security 17), p. 19, Aug. 2017.

[55]. Crossler, F. Belanger, and D. Ormond, "The quest for complete security: An empirical analysis of users multi-layered protection from security threats," Information Systems Frontiers, Apr. 2017. [Online].