International Journal of Multidisciplinary Global Research
Vol. 02 (01), Jan-Mar 2025, pp. 01-09.
Online ISSN: 3107-4197

Advanced In-Depth Literature Review on Intrusion Detection Systems

NASRUDDIN ANNAPURI*

Independent Researcher, Department of Artificial Intelligence and Data Science (AIDS)
Parul University, Vadodara, Gujarat, India
nasruddinannapuri@gmail.com

KARTHIKEYA ANNAM

Independent Researcher, Department of Artificial Intelligence and Data Science (AIDS)
Parul University, Vadodara, Gujarat, India
karthikeyaannam@gmail.com

MUJTABA ASHRAF QURESHI

Department of Artificial Intelligence and Data Science (AIDS)
Parul University, Vadodara, Gujarat, India
mujtaba.qureshi36782@paruluniversity.ac.in

Abstract

As cyber protective systems become more advanced, they rely on constant intrusion monitoring and tackling access, attacks, and associated renewing dangers to disguise malicious activities within the network. Modern cybersecurity designs would not be complete without intrusion detection systems (IDS), which monitor networks for suspicious activity and new threats. Three groups are examined in this work regarding IDS transformation: signature-based detection, anomaly-based detection, and hybrid models involving AI and ML. Federalized learning, blockchain integration, and zero-trust architectures are some of the new areas of research that are being thoroughly discussed. Other topics covered include the underlying methodologies, the integration of advanced AI/ML techniques, technical challenges like scalability and encrypted traffic analysis, and emerging trends in the field. The review draws upon diverse sources from IEEE, ACM, Springer, IOPscience, and ScienceDirect.

Keywords: Anomaly- Based Detection, Blockchain, Cybersecurity, Federated Learning, Hybrid IDS, Intrusion Detection Systems, Machine Learning, Zero Trust.

1. Introduction

Cybersecurity, in the present-day generation, has emerged as a worldwide vital, pushed by the essential need to shield systems from undesirable, unauthorized, and unexpected interference. Those interferences can range from facts breaches and facts robbery to threats that undermine the integrity and functionality of systems. The

*Corresponding Author.

rapid evolution of cyber threats, driven by increasing network complexity and sophisticated attack vectors, demands equally advanced security measures. Modern threats have rendered outdated protections like antivirus and firewall programs useless. The secondary line of defence against prospective intrusions is the intrusion detection system (IDS), which keeps a constant eye on system activities and network traffic to spot suspicious activity. When it detects suspicious activity on a network or device, an Intrusion Detection System (IDS) immediately notifies the appropriate parties and keeps tabs on any suspicious transactions that may be taking place. This review provides:

- An overview of the evolution of IDS from static, rule-based approaches to dynamic, AI/ML-enhanced models.
- A detailed discussion of various IDS methodologies: Three types of systems are signature-based, anomaly-based, and hybrid.
- An explanation of how the IDS works in general.
- Detailed examination of the supervised, unsupervised, and deep learning methods used by AI and ML in intrusion detection systems.
- High false positive rates, problems with scalability, and dealing with encrypted traffic are some of the present difficulties that this article investigates.
- A look at emerging trends and future directions, including federated learning, blockchain-based security, and zero-trust architectures.

Numerous studies have highlighted the necessity for resilient and adaptive IDS solutions (Amarudin et al., 2020; Issa et al., 2024; Chauhan et al., 2020).

2. Taxonomy of IDS Approaches

IDS approaches are broadly categorized by their detection methods and their deployment location. There are three main kinds of intrusion detection systems, distinguished by the way they identify intrusions. There are three types of systems: hybrid, signature-based, and anomaly-based. Based on their placement within the network, intrusion detection systems can be classified into three distinct types. The three main types are those that rely on networks, hosts, and application protocols.

2.1. *Signature-Based IDS*

Mechanism: Signature-based IDS scans network traffic or system logs, and matches against a repository of known attack signatures. Alerts are raised if a match is found. This works well for established threats. Ordinary updates are needed to detect new threats, but unknown assaults without signatures can bypass this gadget.

Advantages:

- **High Accurate**: Detects known attacks with minimal false alerts as long as the signatures are up to date.
- **Productivity**: Does not use a vast amount of processing power.

Limitations:

- **Novel Attacks**: Attacks that have not yet been catalogued are wholly undetectable.
- **Database Management**: Must regularly refresh the signature database.

Example: Snort and Suricata are very popular in the industry (Amarudin et al., 2020).

2.2. *Anomaly-Based IDS*

Mechanism: Some use statistical analysis and machine learning to construct a model of perfectly typical behavior, while others are based on anomalies. Any positive deviation from this baseline is flagged as an intrusion attempt. By keeping tabs on device activity and labelling it as normal or unusual, an anomaly-based intrusion detection system can identify both network and computer intrusions and misuse.

Advantages:

- **New Threat Detection**: They can try to detect something very new they never saw before or zero-day attacks.
- **Adaptability**: A network can adapt over time. Limitations:

Limitations:

- **False Alarms**: They tend to raise false alarms with no real threat just because some behavior is out of the ordinary.
- **Highly Demanding**: Needs high computation resources and precise calibration of the model automation.

Example: Autoencoders have been shown to detect anomalies in network traffic very effectively (Ashiku & Dagli, 2021).

2.3. *Hybrid IDS*

Mechanism: Hybrid IDS unites both signature and ''anomaly outlier" based IDS to exploit their respective opportunities. These systems cross-validate alerts, attempting to reduce false positives while improving overall detection accuracy. When compared to other intrusion detection methods, the hybrid system performs better.

Advantages:

- **Complementary Features**: Combines the accurate detection of known threats with the ability to identify unknown attacks.
- **Reduce Errors in Validation**: Drives the rate of one or both false positive and negative results down.

Limitations:

- **Simplicity**: Has one approach but this one is a costly approach because it combines different comprehensive approaches which build up the system's complexity.
- **Processing Requirements**: Needs more powerful processors with precise system structuring to enable these approaches.

Reference: Extensive research on hybrid IDS is available in the literature (Khraisat et al., 2019).

2.4. *Network-Based IDS*

Mechanism: Network-based Intrusion Detection Systems, or NIDS, perceive attacks by analyzing entire network traffic including individual packets that move over the network against known attack signatures, or using heuristics and statistical-based anomaly detection systems.

Advantages:

- Wider Coverage: Monitors multiple devices in a network without requiring agents on individual hosts.
- Scalability: Can handle large-scale network traffic and provide centralized security monitoring.

Limitations:

- Encrypted Traffic: Mismanaged threat detection for encrypted communications.
- High False Positives: Requires extensive fine-tuning or will generate alerts for benign activities.

Example: Known to most random users and computer saves is "Snort", an open-source product for NIDS intrusion detection in the system (Arqane et al., 2021).

2.5. *Host-Based IDS*

Mechanism: Tailored for workstations and servers, these host-based computers monitor file activities, system-generated logs, and process events, looking for irregularities that would indicate settlement and consequently an invasion. They accomplish this by tracking system calls and the behavior of programs hosted by a given system and flagging any anomalies above a predefined threshold. If the analytical machine documents have been edited or deleted, an alert is sent to the administrator to analyze.

Advantages:

- Deep System Visibility: Can detect internal attacks and unauthorized access to files.
- Customizable: Tailored to individual host environments for better precision.

Limitations:

- **Resource Intensive**: This can consume significant CPU and memory on the host.
- **Limited Scope**: It only tracks the host where it is installed, ignoring the other network dangers.

Example: Famous among open-source HIDSs, OSSEC analyses logs and checks files for integrity (Chauhan et al., 2020).

2.6. *Application Protocol-Based IDS (APIDS):*

Mechanism: Intrusion detection systems that focus on application-layer protocols (APIDS) include those that keep an eye out for malicious behavior by analyzing and monitoring protocols such as HTTP, FTP, DNS, and SMTP. They are specialized in finding anomalies, or known attack signatures, in a particular application.

Advantages:

- **Deep Packet Inspection**: It analyzes application-layer data profoundly.
- **Against Web Attacks**: Helps to detect SQL injections, cross-site scripting (XSS), and other application-layer attacks.

Limitations:

- **Complex Configuration**: Needs a good understanding of application protocols to achieve adequate effectiveness.
- **Performance Overhead**: The deep inspection may slow down the network communications.

Example: ModSecurity is a well-known example of a Web Application Firewall (WAF) that works as an APIDS by preventing web attacks (Nisa et al., 2022).

3. How a detection System is breached?

In a nutshell, the number one priority of IDS is to recognize irregularities without harm being done to the device and network by cybercriminals. Automated processes or tools use a data repository of attacks or a dissimilar as to how the network functions for "normal" data retrieval to seek out differences triggered by deviations.

The system then raises red flags on the identified anomalies utilizing the deviation detection systems for supplementary review and analysis inspections in the networking and application layers. The typical foundation of an IDS is comprised of up to three components which are: Dee Tweeter 2010 Security sensors that process activities, both user and computer, over the network to initiate a security event.

By monitoring the events, notifications, or operational response to the reported alerts, the console manages that information with the intent to issue notifications. All events that concern security incidents together with their cheers, stoops, and alerts are documented by the detection engine and put in a special archive.

Additionally, sign detection, pattern detection, protocol analysis, and statistical analysis provide malicious traffic with four unique pinpointing approaches that an IDS enlisted.

3.1. *Signatures*

By utilizing already known patterns of attacks, one can detect the IDS signature against the network data signature of packet material.

3.2. *Anomalies*

Nowadays, machine learning techniques are observed in the more advanced systems to identify malfunctioning or abnormal behaviour in the flow of traffic within the network or datagram. The super-vised ML algorithm autonomously identifies patterns and is developed through frequent collaboration with network events.

3.3. *Unauthorized Access*

Using an Access Control List(Create ACL), security access boundaries are configured and observed.

4. The Influence of AI and Machine Learning on IDS

Automatic feature extraction and adaptive learning procedures are two ways in which IDS has been transformed by the use of AI and ML technology.

4.1. *Techniques of Supervised Learning:*

Approach: Supervised learning requires a pre-labelled dataset where traffic is differentiated between normal and malicious so that models can be trained on it.
Key Algorithms:

- **Support Vector Machines (SVM)**: Impressive results on challenges requiring binary categorization.
- **Decision Trees and Random Forests**: High accuracy as well as great interpretability.

Strengths:

- Very high accuracy of classification with complete datasets.
- Some models provide great transparency about how decisions are made.
- Limitations:
- Highly reliant on quality-labeled training sets. Not proficient in generalizing novel attacks.

Example: SVM and Random Forest models have been successfully implemented in IDS (Eriza & Survadi, 2021).

4.2. *Techniques of Unsupervised and Deep Learning*

Approach: Unsupervised methods such as clustering or even autoencoders capture data patterns without pre-assigned labels. CNNs and RNNs are deep learning models that also feature automatic extraction of features from the direct network data feed.
Strengths:

- Used to detect zero-day attacks, these models are efficient.
- Elementary feature extraction is automatic, necessitating little to no preliminary processing.

Limitations:

- Computationally intensive.
- Without sufficient masking, there is a chance of overfitting.

4.3. *Detailed Mechanisms*

1) **Collection of Data and Feature Selection**: IDS gathers information like network traffic, logs, and user behavior. Important features are extracted with techniques like Principal Component Analysis (PCA) and autoencoders.
2) **Training and Evaluation of Models**: Supervised models learn from labelled data; unsupervised models establish normal behavior patterns and learn from them. Popular measures of performance include recall, accuracy, precision, and the F1 score.
3) **Adaptation in Real Time**: Online learning or adversarial training offers methods where models can continuously respond to new threats.
4) **Federated Learning**: Newer approaches allow multiple organizations to create a federated model without exposing the data, enabling collaborative training while maintaining confidentiality.

Case Study: The real-time detection functionality of CNN-based IDS has performed exquisitely well (Eriza & Survadi, 2021).

5. Challenges in Modern IDS Implementation

Modern-day systems have made great strides yet as noted in the above example there are challenges as well.

5.1. *Façade CT False Positives and False Negatives*

An ever-present difficulty is striking a balance between being too sensitive and being too particular. Systems that are too sensitive emit too many false positives, while those that are not sensitive enough may overlook true threats. Adaptive thresholding and ensemble techniques for hybrid systems are important in these types of systems (Khraisat et al., 2019).

5.2. *Related to System Size and Response Time*

In the present day, networks generate an enormous amount of data, this pushes for IDS systems that function in real-time. Many solutions attempt to solve scaling and latency such as distributed processing frameworks, edge computing, and cloud computing (Arqane et al., 2021).

5.3. *Approaches to Process Encrypted Traffic*

The ubiquitous deployment of encryption methodologies renders deep packet inspection less effective. Some researchers are focusing on threat detection approaches such as metadata analysis and selective decryption that do not compromise user privacy (Ashiku & Dagli, 2021).

6. New Directions and Emerging Developments

With the increase in scale and sophistication of modernized cyber threats towards IDS, there is a need to integrate novel technologies that can improve their detection accuracy, efficiency, and veracity. Academics have been paying attention to innovative ideas in the field of distributed intrusion detection systems (IDS) for a while now, such as distributed federated learning (FIDL), threat intelligence sharing, and blockchain-based secure logging. These advancements tackle significant issues like privacy-preserving collaborative machine learning, secure alert border control, and mitigation of insider and outsider threat adaptive security postures. In this part of the paper, we will analyze those recent changes in IDS development direction by presenting all possible advantages, obstacles, and gaps for further exploration.

6.1. *Using Federated Learning for a Distributed Intrusion Detection System (IDS) Learning for Distributed IDS*

(Khraisat et al., 2019) state that federated learning improves detection capabilities while preserving privacy by allowing collaborative model training across different entities without sharing raw data.

6.2. *Integration with Blockchain*

Blockchain enables powerful decentralization due to its immutable nature, allowing for safe and reliable IDS alert logging and threat intelligence sharing.

6.3. *Architectures Based on Zero Trust*

The constant need for verification at every network access request drastically decreases external and internal threats. The integration of zero-trust principles into IDS is perceived as a hopeful solution for future security systems.

7. Conclusion

Although the integration of AI and ML into biometrics techniques offers improved threat detection, problems like false negatives, scalability, and encrypted traffic assessment continue to pose serious difficulties. Improving the limitations uses multipronged approaches: Federated Learning, Blockchain, and Zero Trust Architectures sharpen decentralization, privacy, and continuous validation all the while enhancing detection accuracy. Instead of focusing on accuracy, increasing computational efficiency and responsiveness can better aid in real-time threat neutralization. The more complex the cyber threats, the more essential it becomes to enforce ID integration with AI, FL, blockchain, and zero trust framework to form robust and flexible security infrastructures that can withstand cyber threats.

8. References

Amarudin, R., Ferdiana, R., & Widyawan. (2020). A systematic literature review of intrusion detection system for network security: Research trends, datasets and methods. 2020 4th International Conference on Informatics and Computational Sciences (ICICoS) (pp. 1–6). IEEE. https://doi.org/10.1109/ICICoS51170.2020.9299068.

Arqane, A., Boutkhoum, O., Boukhriss, H., & El Moutaouakkil, A. (2021). NISS '21: Proceedings of the 4th International Conference on Networking, Information Systems & Security (Article No. 7, pp. 1–6). ACM. https://doi.org/10.1145/3454127.345657

Ashiku, L., & Dagli, C. (2021). Network intrusion detection system using deep learning. Procedia Computer Science, 185, 239–247. https://doi.org/10.1016/j.procs.2021.05.025

Chauhan, A., Singh, R., & Jain, P. (2020). [Title not provided]. Journal of Physics: Conference Series, 1518(1), 012040. https://doi.org/10.1088/1742-6596/1518/1/012040

Eriza, A. A., & Survadi, M. T. (2021). Literature review of machine learning models on intrusion detection for Internet of Things attacks. 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1–5). IEEE. https://doi.org/10.1109/ICECET52533.2021.9698760

Issa, M. M., Aljanabi, M., & Muhialdeen, H. M. (2024). Systematic literature review on intrusion detection systems: Research trends, algorithms, methods, datasets, and limitations. Journal of Intelligent Systems, 33(1), 20230248. https://doi.org/10.1515/jisys-2023-0248

Khair Ul Nisa, Qureshi, M. A., & Ahmad, A. (2022). High-level security approach in wireless sensor network using cluster-based dynamic keying technique. Journal of Algebraic Statistics, 13(2), 1523–1532.

Khraisat, A., Gondal, I., Vamplew, P., & Kamruzzaman, J. (2019). Survey of intrusion detection systems: Techniques, datasets and challenges. Cybersecurity, 2(1), 20. https://doi.org/10.1186/s42400-019-0038-7

Qureshi, M. A., & Sheikh, M. I. (2021). Operational business prediction system using machine learning algorithm. Design Engineering, 12511–12524.

Sheikh, M. I., & Qureshi, M. A. (2021). Effectiveness of modern technology in decline the contagious clout of COVID19. Design Engineering, 12920–12932.