Critical Review of Survey Paper: Deep Learning-based Intrusion Detection Systems

Paper Reviewed: Deep Learning-based Intrusion Detection Systems: A Survey (arXiv, 2025)

Student Name: Yasir Alalawi

Student ID: 447540166

This review is about the survey *"Deep Learning-based Intrusion Detection Systems: A Survey (2025)"*, which shows how deep learning methods are used in modern intrusion detection systems (IDS). The survey organizes existing research through a structured workflow perspective, emphasizing how different components of IDS development connect in practice. In the following sections, I discuss the survey's scope, methodological coverage, key strengths, limitations, and its suggested future directions, following the structure required for this assignment.

1. Overview

1.1 Domain Coverage

The surveyed work falls within the broader field of cybersecurity and focuses specifically on intrusion detection systems (IDS). IDS are security mechanisms intended to detect malicious behavior, unauthorized access attempts, or unusual system activity. While the survey primarily concentrates on network-based intrusion detection systems (NIDS), it also acknowledges host-based IDS (HIDS) and hybrid configurations.

One of the survey's parts is how it frames IDS research as a multi stage process ,not one simpl comparison of model architectures. The authors organize the field into the following phases: process as consisting of raw data, collection, storage, parsing, summarization, detection, and investigation.. This framing reflects an effort to connect deep learning research with the practical structure of operational security systems.

From a technical point of view, intrusion detection is not a simple problem for deep learning. Network traffic contains many different types of information, such as IP addresses, ports, protocols, timestamps, and other metadata. In addition, malicious traffic is much rarer than normal traffic, which creates class imbalance. Traffic patterns also change over time, and detection systems often need to work in near real time. Finally, attackers actively try to avoid detection, which makes the problem more complex than standard classification tasks. Because of these factors, applying deep learning to IDS is challenging in practice.

1.2 Types of Methods Discussed

The survey also categorizes deep learning-based IDS methods based on how they learn and which neural network architectures they use. This makes it easier to understand how different approaches are applied to intrusion detection problems.

In supervised learning, the survey discusses common neural network models such as CNNs, RNNs, LSTMs, and fully connected deep neural networks. CNNs are often used when traffic features are organized in structured formats, while RNNs and LSTMs are useful for analyzing sequences over time. In many studies, these models are trained on processed network flow features rather than raw packet data.

In scenarios where labeled attack data is rare, the survey examines unsupervised and semi-supervised techniques such as autoencoders, variational autoencoders (VAEs), and deep belief networks (DBNs). These approaches are often used for anomaly detection, especially when labeled attack data is limited. Instead of learning specific attack patterns, the model learns what normal behavior looks like and flags unusual deviations.

1.3 Main Contributions

The survey makes several important contributions. First, it introduces a workflow-based structure that organizes IDS research into stages such as data management and detection. This helps readers understand how deep learning fits into the overall intrusion detection process instead of looking at models in isolation.

Second, the survey reviews widely used IDS datasets and explains their strengths and limitations. Since model performance depends heavily on the dataset being used, this discussion helps put reported results into context.

Third, the survey points out ongoing research challenges such as scalability, robustness against attacks, and real-world deployment issues. Rather than presenting deep learning as a complete solution, the authors acknowledge that there are still practical limitations.

2. Critical Evaluation

2.1 Coverage of the Field

Overall, the survey provides a clear and organized overview of how deep learning is applied in intrusion detection. It groups the research using its workflow structure and also categorizes methods into supervised, unsupervised, hybrid, and emerging approaches. This makes the field easier to understand and follow.

However, the discussion is more detailed for network-based intrusion detection than for other IDS settings. Host-based systems and some deployment environments receive less attention. As a result, the coverage is stronger in certain areas than others.

In addition, although newer architectures such as transformer-based models are mentioned, they are not discussed in depth. Since intrusion detection often involves analyzing sequences of events, a deeper discussion of these sequence-oriented models could have strengthened the survey.

2.2 Underexplored or Missing Areas

Several practical aspects are not discussed in much detail.

First, the survey focuses mainly on detection accuracy but gives less attention to operational constraints such as inference speed, memory usage, and computational cost. In real-world environments, IDS systems must meet strict performance requirements. Without clearer discussion of these factors, it is difficult to evaluate how practical some of the proposed models are.

Second, adversarial robustness could have been explored further. Intrusion detection operates in settings where attackers actively try to bypass detection. Although the survey mentions evasion and poisoning attacks, it does not deeply examine defensive strategies or standardized robustness evaluation methods.

Third, concept drift is only briefly addressed. Network traffic changes over time, but many reviewed approaches assume relatively stable data distributions. While adaptability is suggested as a future direction, the survey does not thoroughly analyze how existing models perform under distribution shifts.

Finally, explainability receives limited discussion. In practice, analysts need to understand why an alert was generated. High accuracy alone is not enough if model decisions are difficult to interpret. A deeper discussion of explainable approaches tailored to IDS would have strengthened the survey's practical relevance.

2.3 Benchmarks and Datasets

The survey reviews several commonly used intrusion detection datasets, including KDDCup99, NSL-KDD, UNSW-NB15, CICIDS, and TON_IoT. This overview makes it easier to compare results across studies and understand how models are evaluated.

However, many of these datasets have known limitations. Some are based on outdated traffic, some rely on synthetic data, and many suffer from strong class imbalance. While the survey mentions these issues, the dependence on static benchmark datasets remains a concern. Good performance on these controlled datasets does not always mean the model will perform well in real enterprise environments.

Overall, the need for more realistic and continuously updated datasets remains an important challenge in deep learning-based intrusion detection.

Table 1 summarizes the main IDS datasets discussed in the survey and highlights their key limitations.

Table 1

Dataset	Year	Noted Limitations	Strengths
KDDCup99	1999	Outdated traffic,duplicates	Historical benchmark
NSL-KDD	2009	Still limited realism	Improved over KDD99
UNSW-NB15	2015	Partial modern coverage	More recent attack types
CICIDS	2017+	Controlled environment	Diverse attack scenarios
TON_IoT	2020	Still synthetic elements	IoT-focused modern data

2.4 Evaluation Metrics and Practical Performance

The survey reviews common evaluation metrics such as accuracy, precision, recall, F1-score, and false positive rate. These metrics are widely used and make it easier to compare results across different studies.

However, intrusion detection is not only a classification problem it is an operational security task. Even a small false positive rate can create a large number of alerts in high traffic networks. In practice, factors such as detection speed, computational cost, and robustness against adaptive attackers are also important, but they are not always reflected in standard evaluation metrics. A stronger focus on deployment oriented measures, such as time to detection and cost aware evaluation, would better capture real-world performance.

3. Future Work and Research Directions

3.1 Open Research Challenges

Although deep learning-based IDS has improved over the years, several challenges remain.

Generalization across environments:
Many models perform well on specific benchmark datasets but struggle when tested on different networks. Improving cross environment performance, for example through transfer learning or domain adaptation, is still a problem.

Data quality and realism:
Public datasets are often outdated or partially synthetic, and collecting accurate labeled intrusion data from real operational environments is difficult. More realistic and continuously updated datasets are needed, along with learning approaches that can handle limited labeled data.

Robustness to adversarial manipulation:
Deep learning models can be affected by evasion and poisoning attacks. Developing stronger evaluation methods and training strategies to improve robustness remains an important research direction.

Adaptability and explainability:
Network traffic changes over time, which can cause model performance to degrade. IDS models need to adapt to concept drift while still providing explanations that analysts can understand and trust.

3.2 Proposed Research Project

One possible research direction is to design an adaptive and explainable deep learning-based IDS for dynamic network environments.

The system could include a continual learning component that updates the model over time, a drift detection mechanism to identify changes in traffic patterns, and an explainability layer to help analysts understand why alerts are generated. Instead of focusing only on benchmark accuracy, the evaluation would emphasize long-term stability, robustness, and practical usability.

This proposal builds directly on the limitations highlighted in the survey and aims to move IDS research closer to real-world deployment needs.

4. Conclusion

This review analyzed the survey *"Deep Learning-based Intrusion Detection Systems: A Survey"* and discussed its main contributions and limitations. The survey provides a clear overview of deep learning methods, commonly used datasets, and ongoing research challenges in intrusion detection.

At the same time, several practical issues remain, particularly in areas such as deployment constraints, adversarial robustness, concept drift, and explainability. While deep learning has shown strong experimental results, real-world reliability requires more attention to robustness, adaptability, and operational needs.

Overall, progress in deep learning-based IDS will depend not only on new model architectures but also on improving how these systems perform in realistic cybersecurity environments.

References

1. Author(s). Deep Learning-based Intrusion Detection Systems: A Survey. arXiv, 2025.
2. Buczak, A. L., & Guven, E. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE Communications Surveys & Tutorials.
3. Sommer, R., & Paxson, V. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. IEEE Symposium on Security and Privacy.