International Journal of Research and Review in Applied Science, Humanities, and Technology
Vol 3, Issue 1 January-March 2026 ISSN: 3048-975X
https://ijrasht.com/

Intrusion Detection Systems using Supervised Machine Learning Techniques: A survey

Chetan Negi, Computer Science & Engineering, SDBCT, Indore, MP, India, chetannegi1755@gmail.com
Pooja Hardiya, Computer Science & Engineering, SDBCT, Indore, MP, India, poojahardiyacs@gmail.com

Abstract— This study examines intrusion detection from the standpoint of supervised machine learning, with the objective of organizing existing research into a comprehensive taxonomy that links intrusion detection systems with supervised learning techniques. To achieve this, the paper presents an in-depth discussion of the fundamental perceptions of intrusion detection systems, normally used supervised machine learning algorithms, and various categories of cybersecurity attacks. Subsequently, prior research efforts that apply supervised learning methodologies to intrusion detection are systematically reviewed and analyzed. Based on this review, a taxonomy is developed to organise and compare remaining approaches. The findings derived from this taxonomy indicate that supervised learning models demonstrate strong and encouraging classification performance when estimated on four widely used intrusion detection datasets: KDD'99, NSL-KDD, CICIDS2017, and UNSW-NB15. Additionally, the study highlights the critical role of feature selection, which is often necessary to improve detection accuracy and reduce computational complexity. Data imbalance is also identified as a significant challenge in intrusion detection datasets, and the use of appropriate sampling techniques is shown to effectively mitigate this issue. Finally, the analysis suggests that for achieving optimal performance on important intrusion detection datasets, deep learning-based supervised methods are particularly well suited.

Keywords— intrusion detection systems, cyber security, supervised machine learning algorithms

Introduction

In the contemporary landscape of pervasive digital interconnectedness, the escalating sophistication and volume of cyber threats necessitate robust defense mechanisms, rendering Intrusion Detection Systems indispensable for maintaining network security and data integrity[1]. Traditional Intrusion Detection Systems, often reliant on signature-based methods, are increasingly inadequate against novel and adaptive attack vectors, thereby highlighting the imperative for more intelligent and flexible solutions [2]. This has spurred extensive research into the application of advanced machine learning and deep learning algorithms to enhance IDS capabilities, enabling the identification of both known and unknown threats [3]. This systematic literature review aims to synthesize the current state-of-the-art in intrusion detection, focusing on supervised machine learning techniques that have significantly transformed these systems by enhancing security, adaptability, and scalability [4]. Specifically, this paper will delve into various supervised learning paradigms, such as decision trees, artificial neural networks, and support vector machines, critically evaluating their respective strengths and limitations in discerning malicious activities from legitimate network traffic [5]. The structured taxonomy of this review further categorizes these techniques into broad classifications, encompassing machine learning, deep learning, optimization algorithms, and the datasets employed, providing a comprehensive overview of advancements within each domain [6]. Findings from this comprehensive analysis indicate that Support Vector Machines, Convolutional Neural Networks, decision trees, and Genetic Algorithms are among the leading techniques for achieving high classification performance in IDS [7]. For instance, Support Vector Machine algorithms have demonstrated superior accuracy and precision in detecting network intrusions across benchmark datasets like KDD and CIC-IDS2017 [8].

Whenever an input event displays patterns that are like those of known bad invasions, the system labels these occurrences as malicious and flags them for further investigation. With a low rate of false positives, these systems have the potential to be effective in recognising known harmful attacks and flagging them for further investigation. One downside of these systems, on the other hand, is that they are incapable of detecting new attacks [9]. When observed events behave in a manner that is considerably different from previously established known good patterns, alarms are generated in anomaly-based systems, and the system is said to be in trouble. The advantage of these systems over signature-based systems is that, in contrast to signature-based systems, they can identify new and changing threats, while signature-based systems are not. Anything that departs from what is considered typical, normalised, or predicted in a particular scenario is referred to as an anomaly, according to the definition. Anomalies are deviations from the expected behaviour of a system that occur only seldom and are characterised as anomalous behaviour. Any occurrence or series of events that deviates from a specified set of usual behaviours must first be identified by an anomaly detection system before the system can recognise the occurrence or series of events as abnormal [10]. It is important to realise that not all aberrations in nature are malicious in their intent. As defined by the term "anomaly," anomalies are simply deviations from expected normal behaviour, which is exactly what they are. As soon as a certain occurrence or pattern is discovered as an anomaly, it may be classed as either benign or malevolent in nature, depending on the circumstances around it [11]. When it comes to anomaly-based systems, one of the most challenging issues is the issue of creating a high rate of false positives as well as a high rate of false negatives, which is also one of the most challenging issues in computer science.
Intrusion Detection System

This system detects intrusions into a network. IDS (Intrusion Detection System) is a PC-based information framework designed to gather party information about noxious actions in a movement of centralised IT resources, separate information, and reply to a predetermined security game plan. A movement of exercises that attempt to deal with the dependability, grouping, or availability of framework resources may be described as an interruption in the flow of work. Interruption may manifest itself in a variety of ways, including noxious activities, unapproved individuals, and those who are already authorised but are striving to gain further benefits. Figure 1 illustrates how an interruption finding framework may be thoroughly constructed within the constraints of two boundaries [1]:

- Analysis approach may define misuse and anomaly IDS.
- Source of information: Host-based IDS vs. Network-based IDS

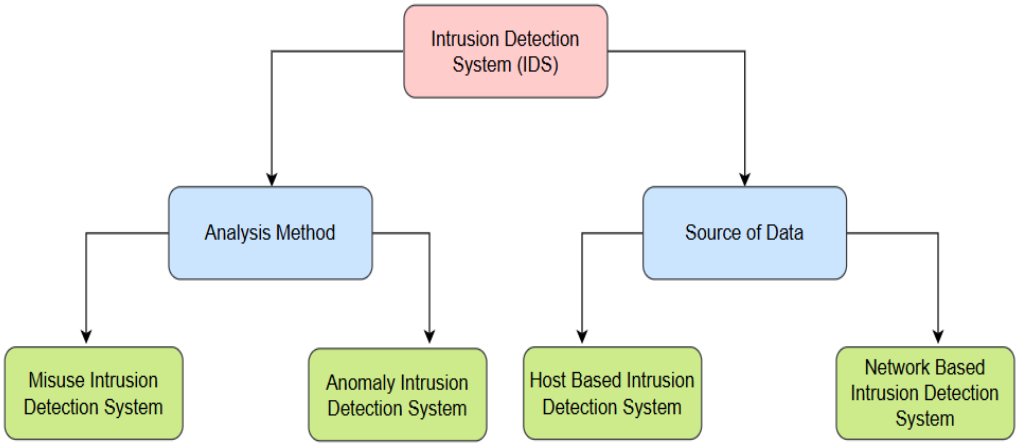

Figure 1: Taxonomy of Intrusion Detection System

Artificial Intelligence: Computer scientists set out in the 1950s to discover whether or not computers could "think" in the same way that human did. This was the beginning of the area of Artificial Intelligence (AI). Marvin Minksy of MIT defines artificial intelligence as "the study of having computers perform tasks that would need intelligence if done manually [1]."

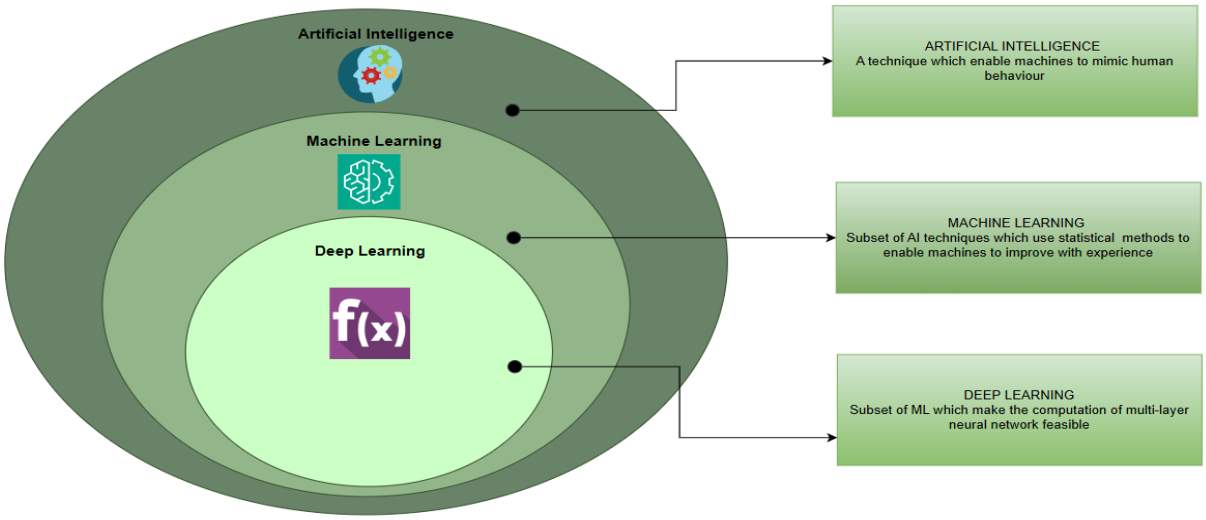

Figure 2: Artificial Intelligence with Machine Learning & Deep Learning

Machine Learning: General-purpose computers may learn and become creative, according to a research published in the journal "Computing Machinery and Intelligence" [3]. Computers can learn from data, rather than having people manually write rules, therefore this arose the question of whether or not computers can learn to accomplish a certain task on their own. These questions led to the creation of the subfield of machine learning. Machine learning algorithms are algorithms that learn from and adapt to the data they encounter. A computer software may learn what output to provide implicitly based on examples and data by using machine learning methods instead of explicitly

programming and guiding it directly. Using examples and data, a computer may be taught how to make decisions and carry out tasks on new inputs it has never seen before [4].

Deep Learning: As previously stated, the fundamental technologies and techniques in deep learning are based on the use of neural network topologies, which are composed of many layers of neurons, to accomplish their goals. After providing some basic information on neural networks, we will explain the differences that exist amongst deep neural network designs [5] in the following section.

2. Literature Review

Numerous studies have corroborated the efficacy of these machine learning approaches, showcasing their ability to identify anomalous network behavior indicative of cyber-attacks [8]. For example, one study achieved a 90% accuracy rate in anomaly identification using system-based properties with Support Vector Machines, suggesting their utility for malware detection at low computational costs [10].

The author of this article discussed and pondered about the current framework in his study. A security invention, interruption recognition, is a security innovation that screens framework in order to keep a strategic distance from malicious activities. Internal Intrusion Detection System and Intrusion Detection System that employs various computations for the framework to function are discussed in detail in this article, which also includes a diagram. Using the concept of interruption recognition, strategies for digital inquiry are developed that use information mining techniques. When compared to previous IDS, the suggested work improves the exactness and identification rate by up to 95 percent, according to the overview. This attribute set, according to the creator, may be used to identify between inside interlopers and their malignant behaviours when designing a new IDS system. It will be a valid IDS that will accurately and continually differentiate the inner gate crasher's and may be used by a few companies to protect their sensitive information. [11]

With the widespread use of personal computers and easy access to the internet on a global scale, the number of methods for attacking a system or a framework has also increased significantly. Entering a building, violating rights, or demanding various people's framework or resources is an illegal practise known as interruption. The primary aim for developing an interruption location framework is to differentiate between attacks on data infrastructure and other types of attacks. It is a security method that attempts to discriminate between various types of attacks. Firewalls are only capable of distinguishing attacks that come from outside of the system, making them ineffective for protecting the system against attacks of any kind. A framework for abuse-based interruption location evaluation was developed in this study. This framework was reviewed in the same way as ALAD, PHAD, LERAD, NETAD and other abnormality-based quantifiable computations. PC organisation is evolving at an accelerated pace these days, and the most astounding tool for a PC organisation is arrangement security[12].

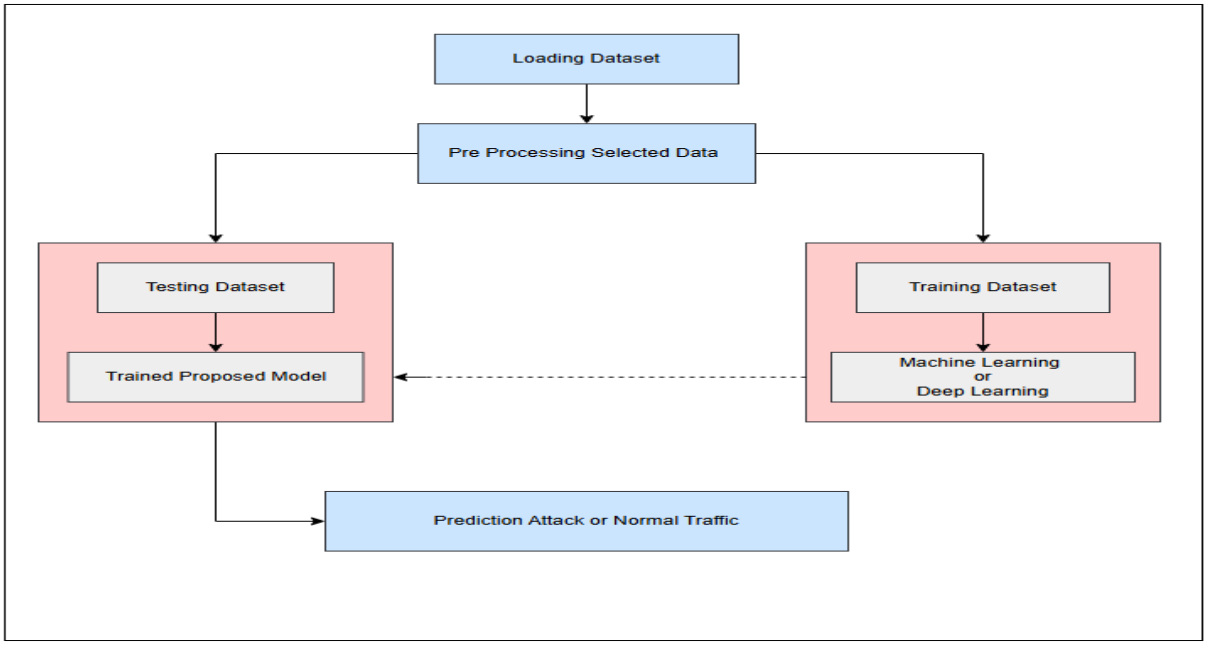

Figure 3: Description of IDS data sets

KDD'99 : It is generated using simulation of normal and attacks traffic in a military environment (US AirForce LAN). It contains nine weeks of simulation in tcp dump files. The dataset is characterized using 41 features related to intrinsic, content, and traffic. Four types of attacks are simulated: DoS, Prob, U2R, and R2L.**NSL-KDD**

UNSW-Nb15: It is a modification to the KDD'99 dataset with solving the problems of redundancy, duplicates, the imbalance of data. It was created using the IXIA Perfect Storm tool to extract normal and attack network traffic based on 100 GB of raw network traffic. It is characterized using 49 feature It consists of around 175 thousand records for training and around 82 thousand records for testing. There are nine types of attacks: Fizzers, Analysis, Backdoor, DoS, Exploit, Generic, Reconnaissance, Shell code, Worm.

CICIDS2017: It was created in an emulated environment in a 5 day period. It contains traffic packet flow and bidirectional flow. 80 features are extracted. Attacks involve: Brute Force FTP, Brute Force SSH, DoS, Heartbleed, Web Attack, Infiltration, Botnet, and DDoS

3. Taxonomy of Intrusion Detection Systems and Supervised Machine Learning Algorithms

This section presents a taxonomy of intrusion detection systems (IDS) and supervised machine learning (ML) approaches derived from the studies reviewed in the previous section. The taxonomy is constructed based on several key attributes summarized in Table 2, including: (1) the dataset utilized, (2) the application of feature selection (Yes/No), (3) the effectiveness of feature selection (Yes/No), (4) the supervised learning algorithms employed, (5) the validation techniques adopted, (6) the best-performing classification algorithm, and (7) the highest reported accuracy and false positive rate (FPR).
To facilitate a clearer understanding of the information reported in Table 2, each dataset is analyzed individually, followed by a consolidated summary of the overall observations.

KDD'99 Dataset: In the case of the KDD'99 dataset, feature selection does not appear to consistently enhance classification performance. Various validation methods have been applied across different studies, making it difficult to identify a sole best-performing algorithm. Nonetheless, both Random Forest and Support Vector Machine classifiers generally exhibit strong presentation and are frequently reported as effective methods for this dataset.

NSL-KDD Dataset: For the NSL-KDD dataset, studies consistently indicate that the use of feature selection leads to noticeable improvements in organization performance. Among the evaluated classifiers, the Random Forest algorithm demonstrates strong and reliable performance across multiple validation strategies. Artificial Neural Networks also achieve promising results on this dataset; however, their evaluation is often limited to a reduced subset of the data, typically around 20%, which restricts broader performance generalization.

CICIDS2017 Dataset: For the CICIDS2017 dataset, no definitive conclusion can be drawn regarding the overall impact of feature selection, although performance improvements are observed in specific cases. A major challenge associated with this dataset is class imbalance, which significantly affects detection accuracy.
Sampling techniques are commonly employed to address this issue. Quadratic Discriminant Analysis emerges as the most effective classification algorithm in the reviewed studies for this dataset.

UNSW-NB15 Dataset: With respect to the UNSW-NB15 dataset, feature selection consistently contributes to improved classification outcomes. Due to the large size and complexity of the dataset, deep learning-based methods, particularly Deep Neural Networks, are frequently accepted and achieve superior performance. These findings suggest that deep learning techniques are well suited for handling large-scale intrusion detection data.

Overall, the following conclusions can be drawn. First, intrusion detection systems that employ machine learning techniques have been lengthily investigated in existing research, sparkly their growing importance in cybersecurity. Second, experimental evaluations conducted on four widely used intrusion detection datasets demonstrate that supervised learning models achieve strong and encouraging classification performance. Third, feature selection plays a vital role in intrusion detection and is often essential for ornamental detection accuracy and dropping model complexity. Fourth, class imbalance remains a significant challenge in intrusion detection datasets; however, the application of suitable sampling techniques can effectively mitigate its negative impact. Finally, when dealing with large-scale intrusion detection data, deep learning-based methods are necessary to achieve high and robust performance.

Table 1: Taxonomy of Supervised Machine Learning–Based Intrusion Detection Systems

Ref.	Dataset	Feature Selection Applied	Feature Selection Useful	Supervised Learning Algorithms	Validation Method	Best Performing Method	Best Reported Results
[10]	NSL-KDD	Yes	Yes	J48, Random Forest	10-fold CV	Random Forest	Accuracy = 99.7%
[11]	NSL-KDD	Yes	Yes	Naïve Bayes, BayesNet, Logistic Regression, Random Forest, J48, Bagging, OneR	Hold-out	Random Forest	Accuracy = 94.0%, FPR = NA
[12]	NSL-KDD	No	NA	SVM, GMM, Random Forest	Hold-out	ANN	Accuracy = 99.0%, FPR = NA
[13]	NSL-KDD (20% sample)	Yes	NA (baseline not reported)	ANN, SVM	Hold-out	ANN	Accuracy = 91.0%, FPR = NA
[14]	KDD'99	Yes	NA (baseline not reported)	PART, ZeroR	10-fold CV	Random Forest	Accuracy = 99.9%, FPR = 0.005%
[15]	KDD'99	Yes	No	SVM	10-fold CV	SVM	Accuracy = 98.7%, FPR = NA
[16]	KDD'99	Yes	No	Bayesian Backpropagation Neural Network (BBNN)	10-fold CV	BBNN	Accuracy = 81.83%, FPR = NA
[17]	CICIDS2017	Yes	Yes	AdaBoost	Hold-out	AdaBoost	Accuracy = 99.0%, FPR = 0.001%

Conclusions

The rapid growth of internet-based services and digital content has contributed to a significant increase in cybercrime. Intrusion Detection Systems represent a fundamental mechanism for identifying and reporting such malicious activities. However, effective anomaly detection remains a challenging task, as it requires accurately distinguishing novel and sophisticated attacks from normal network behavior. These challenges have attracted substantial attention from researchers worldwide, particularly with respect to the application of supervised machine learning techniques for improving intrusion detection performance. In this study, we present a comprehensive review of intrusion detection system classifications, supervised machine learning procedures, and various groupings of cybersecurity attacks. Relevant research efforts are analyzed using four widely adopted intrusion detection datasets: KDD'99, NSL-KDD, CICIDS2017, and UNSW-NB15. Based on this analysis, a taxonomy is developed to methodically categorize existing methods. The resulting classification reveals that intrusion detection using supervised machine learning methods is an actively investigated and rapidly evolving area. Experimental answers across the four datasets indicate that supervised learning algorithms achieve strong and promising classification performance. Moreover, feature selection is shown to be a crucial component for enhancing detection accuracy in many cases. The study also identifies class unevenness as a common challenge in intrusion detection datasets, which can be effectively addressed through suitable sampling techniques. Finally, for large-scale intrusion detection datasets, deep learning-based supervised methods are essential to achieving high and robust detection performance.

References

1. Pinto, A.; Herrera, L.-C.; Donoso, Y.; Gutierrez, J.A. Survey on Intrusion Detection Systems Based on Machine Learning Techniques for the Protection of Critical Infrastructure. Sensors 2023, 23, 2415. https://doi.org/10.3390/s23052415
2. D. Markopoulou and V. Papakonstantinou, "The regulatory framework for the protection of critical infrastructures against cyberthreats: Identifying shortcomings and addressing future challenges: The case of the health sector," Computer Law & Security Review, vol. 41, Art. no. 105502, 2021.
3. G. E. I. Selim, E. E.-D. Hemdan, A. M. Shehata, and N. A. El-Fishawy, "Anomaly events classification and detection system in critical industrial Internet of Things infrastructure using machine learning algorithms," Multimedia Tools and Applications, vol. 80, pp. 12619–12640, 2021.
4. I. Ahmed, M. Anisetti, A. Ahmad, and G. Jeon, "A multilayer deep learning approach for malware classification in 5G-enabled IIoT," IEEE Transactions on Industrial Informatics, vol. 19, no. 2, pp. 1495–1503, 2022.
5. M. A. Ridwan, N. A. M. Radzi, F. Abdullah, and Y. E. Jalil, "Applications of machine learning in networking: A survey of current issues and future challenges," IEEE Access, vol. 9, pp. 52523–52556, 2021.
6. K. Shaukat, S. Luo, V. Varadharajan, I. A. Hameed, and M. Xu, "A survey on machine learning techniques for cyber security in the last decade," IEEE Access, vol. 8, pp. 222310–222354, 2020.
7. L. Kruszka, M. Klósak, and P. Muzolf, Critical Infrastructure Protection: Best Practices and Innovative Methods of Protection, NATO Science for Peace and Security Series D: Information and Communication Security, vol. 52. Amsterdam, The Netherlands: IOS Press, 2019.
8. A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: Techniques, datasets and challenges," Cybersecurity, vol. 2, no. 1, Art. no. 20, 2019.
9. T. T. Nguyen and V. J. Reddi, "Deep reinforcement learning for cyber security," IEEE Transactions on Neural Networks and Learning Systems, pp. 1–17, 2021.
10. J. Graham, R. Olson, and R. Howard, Cyber Security Essentials. Boca Raton, FL, USA: CRC Press, 2011.
11. H.-J. Liao, C.-H. R. Lin, Y.-C. Lin, and K.-Y. Tung, "Intrusion detection system: A comprehensive review," Journal of Network and Computer Applications, vol. 36, no. 1, pp. 16–24, 2013.
12. H. Liu and B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey," Applied Sciences, vol. 9, no. 20, Art. no. 4396, 2019.
13. Y. Hamid, M. Sugumaran, and V. R. Balasaraswathi, "IDS using machine learning—Current state of the art and future directions," Current Journal of Applied Science and Technology, pp. 1–22, 2016.
14. I. H. Witten and E. Frank, Data Mining: Practical Machine Learning Tools and Techniques. San Francisco, CA, USA: Morgan Kaufmann, 2002.
15. A. Sahasrabuddhe, S. Naikade, A. Ramaswamy, B. Sadliwala, and P. Futane, "Survey on intrusion detection system using data mining techniques," International Research Journal of Engineering and Technology, vol. 4, no. 5, pp. 1780–1784, 2017.
16. N. Farnaaz and M. A. Jabbar, "Random forest modeling for network intrusion detection system," Procedia Computer Science, vol. 89, pp. 213–217, 2016.
17. J. Rust, "Using randomization to break the curse of dimensionality," Econometrica, vol. 65, no. 3, pp. 487–516, 1997.
18. University of New Brunswick, "Canadian Institute for Cybersecurity datasets," [Online]. Available: https://www.unb.ca/cic/datasets/index.html . Accessed: Jun. 1, 2021.
19. M. Masdari and H. Khezri, "A survey and taxonomy of fuzzy signature-based intrusion detection systems," Applied Soft Computing, Art. no. 106301, 2020.
20. A. Milenkoski, M. Vieira, S. Kounev, A. Avritzer, and B. D. Payne, "Evaluating computer intrusion detection systems: A survey of common practices," ACM Computing Surveys, vol. 48, no. 1, pp. 1–41, 2015.
21. E. Conrad, S. Misenar, and J. Feldman, CISSP Study Guide. Oxford, U.K.: Newnes, 2012.
22. C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY, USA: Springer, 2006.