The Scientific Temper (2024) Vol. 15 (3): 2719-2724
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2024.15.3.40
https://scientifictemper.com/

RESEARCH ARTICLE

A novel and an effective intrusion detection system using machine learning techniques

Remya Raj B.*, R. Suganya

Abstract

Network environments become more and more diverse with the presence of many different network protocols, services, applications and so on. With this diversity, many different types of attacks appear and target a computer or a network every day. A single type of intrusion detection system (IDSs), which has its own advantages and disadvantages, seems to be insufficient to detect all the attacks. Since we don't know which types of attacks are coming next, the primary difficulty lies in selecting the best IDS at a certain time. In our scenario, we assume that each IDS has its own favorite types of attacks to detect. This paper investigates for intrusion detection system (IDS) and its performance has been evaluated on the normal and abnormal intrusion datasets (KDDCUP99). A new technique of k-NN algorithm using NA (Network Anomaly) rules for intrusion detection systems is experimented. The research work compares the accuracy, detection rate, false alarm rate and accuracy of other attacks under different proportions of normal information. A comparison between Naive Bayes classifier, SVM and NA-kNN for the same training data set and testing data set has been carried out. Experimental results show that for Probe, U2R, and R2L, NA-kNN gives better results. Overall correct count to detect correct attacks is larger in NA-kNN than other classifier algorithms.

Keywords: Intrusion, Security, Supervised, Neighbor, Attack, Rule mining, Machine learning.

Introduction

Network security plays a very important role in today's web-enabled world. In 21st century, network traffic has increased because of enormous growth in online users and their online communication (Singh, K. K. *et al.,* 2004). Number of security attacks has increased with the increase in internet users. The frequency and severity of such attacks have shown a great impact on network performance (Kruegel, C. *et al.,* 2005). Thus, it can be safely argued that despite the variety of existing protection methods described in the literature in recent years, including peripheral protection mechanisms and various authentication and access control techniques, integral protection against intrusions cannot be achieved (Wang, W. *et al*. 2006). One of the way-out to solve this problem is by using an intrusion detection system (IDS). The main function of IDS is distinguishing and predicting normal or abnormal behaviors. An IDS gathers and analyzes information from various sources within computers and networks to identify suspicious activities that attempt to illegally access, manipulate, and disable computer systems. In olden times the concept of intrusion detection appeared in the late 1970s (Jones, A. *et al.,* 2000). Anderson was the first author who had written the first research paper, Computer Security Threat Monitoring and Surveillance (Anderson, J. P. *et al.,* 1980) on Intrusion Detection. Any attempt, successful or unsuccessful, to compromise the confidentiality, integrity and availability of any information resource or the information itself is considered a security threat or an intrusion (Kruegel, C. *et al.,* 2005), (Mukherjee B. *et al.,* 1994).

In 1987, intrusion detection technology became a well-established research area after Denning's seminal paper (Denning, D. E. *et al.,* 1987). Since then, a notable amount of IDPS research has been carried out. Currently, the two basic methods of intrusion detection (analytical method) are signature-based and anomaly-based (Denning, D. E. *et al.,* 1987). The signature-based method (Chebrolu, S. *et al.,* 2005), (Lee, W. *et al.,* 2000), also known as misuse detection (Roech

PG & Research Department of Computer Science, Maruthu Pandiyar College, Affiliated to Bharathidasan University, Vallam, Thanjavur, Tamilnadu, India.

***Corresponding Author:** Remya Raj B., PG & Research Department of Computer Science, Maruthu Pandiyar College, Affiliated to Bharathidasan University, Vallam, Thanjavur, Tamilnadu, India., E-Mail: remyarajphd2021@gmail.com

How to cite this article: Raj, R. B., Suganya, R. (2024). A novel and an effective intrusion detection system using machine learning techniques. The Scientific Temper, **15**(3):2719-2724.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.3.40

Source of support: Nil

Conflict of interest: None.

Received: 15/07/2024 **Accepted:** 20/08/2024 **Published** : 16/09/2024

1999), looks for a specific signature to match, signaling an intrusion. This approach is similar to the way of detecting viruses in many anti-virus applications. A set of patterns of known attacks is necessary to be built in advance for further detection. They can detect many or all known attack patterns, but the weakness of signature-based intrusion detection systems is the incapability of identifying new types of attacks or variations of known attacks.

Machine learning techniques classification is based on the input of training data into three types: supervised, unsupervised and semi-supervised. In unsupervised learning, the input examples are not class labeled. In supervised learning, the labeled examples are used in the training dataset. Expert labeling of the data is very expensive, though there is a huge amount of network and host data available. Semi-supervised learning methods can make use of unlabeled examples in addition to labeled ones. Semi-supervised learning methods require a small quantity of labeled data while still taking advantage of the large quantities of unlabeled data. In the classification phase, the built model or trained classifier is applied to assign the test pattern to one of the pattern classes under consideration of the selected attributes from the training phase. Due to the efficiency of the automatic learning techniques, the machine-learning-based intrusion detection systems (ML-IDSs) allow quickly the attacks while demanding much less manual work. Because of this reason, the approach is becoming more and more important for computer security (Maloof, M. A. 2005), especially when the huge amount of network data that needs to be analyzed by intrusion detection systems is increasing rapidly. However, the ML-IDSs are mostly not being used in practice for information security systems.

The ultimate goal of this research is the definition of an advanced anomaly-driven IDS framework for networks. Such a framework would be capable of detecting new types of attacks. Towards this aim, in the context of this thesis, we explore, propose and evaluate new machine learning approaches and characteristics that enhance network security. To do so, supervised learning, semi-supervised learning and real-time IDS-based detection methods are used in an effort to detect an attack. This allowed us to identify anomaly patterns of activities that deviate from a given pre-defined normal profile.

Also, an important aspect of the current research is to explore and understand how new network security threats are able to shape themselves into attacks that seek to compromise fundamental principles of user security and privacy. This knowledge has been used to create proper security mechanisms for the network using machine learning techniques and further test them thoroughly.

Related Work

A statistical anomaly-based IDS finds out normal network activity like what sort of bandwidth is generally used, what protocols are used, and what ports and devices generally connect to each other and aware the administrator or user when traffic is detected that is anomalous (not normal) (Denning, D. E. *et al.,* 1985), (Ye, N. *et al.,* 2002). It is again categorized into univariate, multivariate and time series models. Univariate model parameters are modeled as independent Gaussian random variables, thus defining an acceptable range of values for every variable. The multivariate model considers the correlation between two or more variables. The time series model uses an interval timer, together with an event counter or resource measure and takes into account the order and inter-arrival times of observations and their values, which are labeled as anomalies if its probability of occurrence is too low at a given time.

Knowledge-based stores information about a subject domain. Information in knowledge-based contains symbolic representations of the expert's rules of judgment in a format that allow the inference engine to perform deduction upon it. The expert system approach is one of the most widely used knowledge-based IDS schemes. Knowledge-based techniques are divided into frame-based model, rule-based models and expert systems. Rule-based is a modified form of the grammar-based production rules. Frame based model localizes an entire body of expected knowledge and actions into a single structure. Expert systems are intended to classify the audit data according to a set of rules involving three steps. First, different attributes and classes are identified from the training data. Second, a set of classification rules, parameters, or procedures are deduced. Third, the audit data are classified accordingly (Denning, D. E. *et al.,* 1985), (Anderson, D. *et al.,* 1995).

Machine learning techniques are based on establishing an explicit or implicit model. A singular characteristic of these schemes is the need for labeled data to train the behavioral model, a procedure that places severe demands on resources. In many cases, the applicability of machine learning principles coincides with that of statistical techniques, although the former is focused on building a model that improves its performance on the basis of previous results. Hence, machine learning for IDS has the ability to change its execution strategy as it acquires new information. This feature could make it desirable to use such schemes for all situations.

Presented a framework of NIDS based on a Naive Bayes algorithm (Bridges, *et al.,* 2000). The framework builds the patterns of the network services over data sets labeled by the services. The framework detects attacks in the datasets using the Naive Bayes classifier algorithm using the built patterns. Compared to the Neural network-based approach, their approach achieves a higher detection rate, is less time-consuming and has a low-cost factor. However, it generates somewhat more false positives. A Naive Bayesian network is a restricted network that has only two layers and assumes

complete independence between the information nodes. This poses a limitation of this research work. In order to alleviate this problem so as to reduce the false positives, active platform or event-based classification may be thought of using a Bayesian network. Researchers have designed several systems dealing with the problem of false alarms in recent years.

Proposed use of Bayesian networks to perform reasoning on complementary security evidence, and thus to potentially reduce false alert rates (Zhai, Y. *et al.,* 2004).

Proposed System

Intrusion is a set of actions aimed at compromising the security of computer and network components in terms of confidentiality, integrity and availability (Heady, R. *et al.,* 1990). An inside or outside agent can do this to gain unauthorized entry and control of the security mechanism. To protect the infrastructure of network systems, intrusion detection systems (IDSs) provide well-established mechanisms that gather and analyze information from various areas within a host or a network to identify possible security breaches.

Intrusion detection functions include:

Monitoring and analyzing user, system, and network activities

The monitoring and analysis function can be represented as:

Monitor(U(*t*),S(*t*),N(*t*))=Analyze(U(*t*),S(*t*),N(*t*))

Configuring systems for generation of reports of possible vulnerabilities

This function involves configuring systems to detect and report potential vulnerabilities based on system parameters and network traffic. It can be expressed as:

GenerateReport(S(*t*),N(*t*))=Report(S(*t*),N(*t*))

Assessing system and file integrity

The function to assess system and file integrity can be denoted as:

AssessIntegrity(S(t))=Integrity(s1(t),s2(t),…,sn(t))

Recognizing patterns of typical attacks

This function involves identifying known attack patterns within network traffic. It can be represented as:

RecognizeAttacks(N(*t*))= PatternRecognition(N(*t*))

Analyzing abnormal activity

This function aims to detect deviations from normal behavior in user activities, system configurations, and network traffic. It can be expressed as:

Analyze-AbnormalActivity(U(*t*),S(*t*),N(*t*)) = Detect Anomalies(U(*t*),S(*t*),N(*t*))

Tracking user policy violations

This function involves monitoring instances where user activities violate established security policies. It can be denoted as:

TrackPolicyViolations(U(*t*))=PolicyViolation(U(*t*))

Methodology

Data Collection Phase

In the data collection phase, the acquisition of reliable and comprehensive datasets is imperative for accurate intrusion detection. We denote the datasets utilized in this study as $\text{Dataset}_{\text{KDDCUP99}}$ and $\text{Dataset}_{\text{DARPA99}}$ referring to the KDDCUP99 and DARPA 99 datasets, respectively. Additionally, to simulate real-time intrusion scenarios, a controlled laboratory environment was established. The laboratory setup facilitated the generation of denial-of-service attacks ($\text{DoS}_{\text{attack}}$) against a victim machine (VM) using tools obtained from the internet ($\text{Tools}_{\text{Internet}}$). It can be expressed as:

$$\text{DoS}_{\text{attack}} = f(\text{Tools}_{\text{Internet}}, \text{VM})$$

An intrusion detection system (IDS) (IDSvictim) was deployed on the victim machine to capture and analyze the incoming traffic (Trafficincoming). The IDS system can be represented as: IDSvictim (Trafficincoming)

Pre-processing procedures were subsequently applied to the collected data (Datacollected) to enhance its quality and suitability for analysis. These procedures represented symbolically as Pre-processing (Datacollected), encompassed data cleaning, normalization, and feature extraction techniques.

Data Pre-processing Phase

This phase is responsible for collecting and providing the log data in the specified form that the feature extraction phase will use. The data preprocessor is, thus, concerned with collecting the data from the desired source and converting it into a format that is comprehensible by the analyzer. KDD Cup99 database has been converted into .arff format. Real-time traffic has been captured with Wireshark. From the packet header, packet size, source address and destination address features are extracted. Data Pre-processing Phase involves data cleaning, normalization, feature extraction, and format conversion.

Cleaning Techniques

To mitigate anomalies and inconsistencies within the dataset, various cleaning techniques were applied. Let X denote the raw dataset, where each row represents a network event and each column corresponds to a specific feature. The cleaning process involved identifying and handling missing values M, outliers O, and duplicate records D. This can be expressed as:

$$\mathbf{X}_{\text{clean}} = \mathbf{X} - (\mathbf{M} \cup \mathbf{O} \cup \mathbf{D})$$

Normalization Methods

Normalization ensures uniformity and comparability across different features. Let X*i* represent the *i*th feature vector of X. Common normalization techniques such as min-max scaling and z-score normalization were applied:

$$\mathbf{X}_{\text{norm}} = \frac{\mathbf{X}_i - min(\mathbf{X}_i)}{max(\mathbf{X}_i) - min(\mathbf{X}_i)}$$

Feature Extraction

Feature extraction aimed to distill relevant information from the raw data. Let F represent the extracted feature matrix obtained through algorithms such as packet parsing and principal component analysis (PCA). The feature extraction process can be formulated as:

$F=f(X)$

Format Conversion

The KDD Cup99 dataset was transformed into the .arff format for compatibility with machine learning algorithms. Let XKDD denote the original dataset and Xarff represent the converted dataset:

$$\mathbf{X}_{\text{arff}} = \text{convert}(\mathbf{X}_{\text{KDD}})$$

Diagnosis and Analysis Phase

The analysis of the intrusion detector phase is the core component that analyzes the traffic patterns to detect attacks. This is a critical component and one of the most researched phases. Here, supervised learning, semi-supervised learning and Real-time detection approaches have been used as intrusion detectors. The capability of the analyzer to detect an attack often determines the strength of the overall IDS.

Post Processing and Expected Alarm class

This phase controls the mechanism to react and determines the best way to respond when the analyzer detects an attack. The system either raises an alert without taking any action against the source or blocks the source for a pre-defined period of time. This action depends upon the security policy that is pre-defined in the IDS. Issues to validate whether the predictions made is correct and related to the actual behavior of IDS implementations is the real challenge. A systematic and complete validation would require that the predictions made by the approach are compared with the behavior of actual IDS implementations. Such an activity would represent an enormous challenge and precisely exemplify the problem that the work attempts to address. It would be required that one or several rather complex environments be built such that IDS can be analyzed under different conditions. However, the most challenging aspect of any such undertaking of validation would be the number and diversity of individual tests to be executed.

Evaluation Metrics

The confusion matrix is a ranking method that can be applied to any kind of classification problem. The size of this matrix depends on the number of distinct classes to be detected. The aim is to compare the actual class labels against the predicted ones. The diagonal represents correct classification. The confusion matrix for intrusion detection is defined as a 2-by-2 matrix since there are only two classes known as intrusion and normal [Ghorbani, A. A. *et al*. 2009], [Dokas, P. *et al*. 2002], [Weiss, S. M. *et al*. 2003]. Thus, the TNs and TPs that represent the correctly predicted cases lie on the matrix diagonal, while the FNs and FPs are on the right and left sides. As a side effect of creating the confusion matrix, all four values are displayed in a way that the relation between them can be easily understood.

Experimental Results

Dataset used for experiments

KDDCup99 dataset is widely used in the experiment of IDS as it provides the basis for comparison of different approaches that require large datasets (KDDCup99 1999), (Stolfo S. J. *et al.,* 2000). In the 1998 DARPA cyber-attack detection evaluation program, an environment (I. S. T. G. MIT Lincoln Lab 2009) was set up to acquire raw TCP/IP dump data for a network by simulating a typical U. S. Air Force LAN. The LAN was operated like a true environment but being blasted with multiple attacks. For each TCP/IP connection, 41 various quantitative (continuous data type) and qualitative (discrete data type) features were extracted among 41 features; 34 features are numeric and seven features are symbolic. The data contains 24 attack types that could be classified into the following four main categories (Table 1):

NA-KNN Algorithm

k value

k value determines the number of nearest neighbors from feature space used to classify a given packet.

Distance Measure Formula

This is a formula to calculate the distance between a given packet and other packets (entities) from feature space. The following distance parameters have been considered to determine the accuracy of IDS by NA-kNN as the Manhattan distance formula. Using these parameters are determined to reduce False Alarm Rate in Intrusion Detection

$$\text{Distance} = \sum_{i \in Features} \sqrt[4]{(X_i - Y_i)^2}$$

Results

Initially, results of the SVM, Naive Bayes classifier and NA-kNN are compared and experimental results show that ADU-kNN gives better results for Probe, U2R and R2L. The results of NA-kNN algorithm for various values of k and distance calculation formulae are compared. After analyzing

Table 1: Dataset specifications

Class	*Class Name*	*No. of Instances*
0	Normal	3973
1	Probe	164
2	DOS	15984
3	U2R	4
4	R2L	39

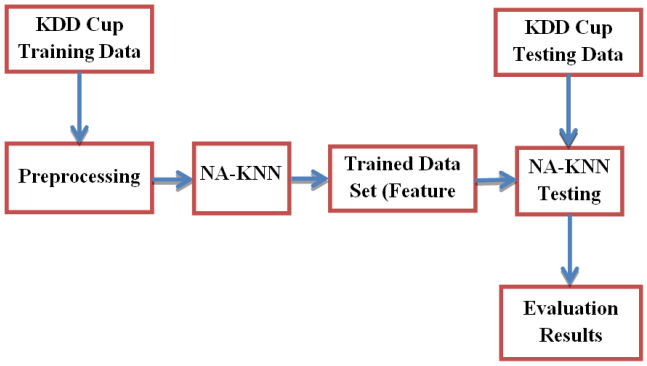

Figure 1: Proposed architecture

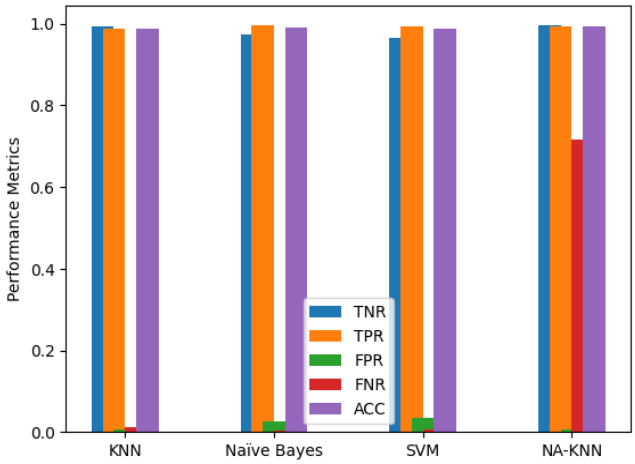

Figure 2: Performance evaluation of proposed systems

different combinatorial models, it is found that the proposed algorithm performs better than other models. It is observed that in NA-kNN model, the increase in the power of distance calculation formulae accuracy increases (Figure 1).

Considering the limit of machines for calculation of the maximum limit of power for a huge number of packets, they are limited up to Manhattan distance. Hence, NA-KNN with Manhattan Distance Formula gives better results. This method gives 99.85% correct result for overall input, for Probe-99.43% detection rate, DOS-99.54% detection rate, U2R-99% detection rate, and R2L-99% detection rate.

Figure 2 and Table 2 illustrates the performance evaluation of intrusion detection algorithms, including KNN, Naïve Bayes, SVM, and NA-KNN, across various metrics such as true negative rate (TNR), true positive rate (TPR), false positive rate (FPR), false negative rate (FNR), and accuracy (ACC). The chart highlights that NA-KNN consistently outperforms other algorithms across most metrics, with notably high TNR and TPR scores and low FPR and FNR scores. Naïve Bayes also demonstrates competitive performance, particularly in terms of accuracy. Conversely, SVM exhibits lower TNR and TPR rates compared to the other algorithms. Overall, the chart emphasizes the effectiveness of NA-KNN, suggesting its suitability for intrusion detection tasks.

Table 2: Performance evaluation of proposed system

Algorithm	*Metrics*	*Probe*	*DOS*	*U2R*	*R2L*
KNN	TNR	0.9924	0.9715	0.9708	0.9685
	TPR	0.9880	0.9955	0.9968	0.9959
	FPR	0.0076	0.0285	0.0292	0.0317
	FNR	0.0120	0.0045	0.0032	0.0041
	ACC	0.9889	0.9855	0.9917	0.9904
Naïve Bayes	TNR	0.9745	0.9546	0.9524	0.9517
	TPR	0.9952	0.9929	0.9932	0.9950
	FPR	0.0255	0.0454	0.0476	0.0483
	FNR	0.0048	0.0071	0.0068	0.0050
	ACC	0.9912	0.9855	0.9852	0.9865
SVM	TNR	0.9655	0.9474	0.9456	0.9441
	TPR	0.9923	0.9871	0.9937	0.9938
	FPR	0.0345	0.0526	0.0544	0.0559
	FNR	0.0077	0.0129	0.0036	0.0062
	ACC	0.9870	0.9859	0.9806	0.9840
NA-KNN	TNR	0.9947	0.9945	0.9909	0.9947
	TPR	0.9945	0.9956	0.9970	0.9969
	FPR	0.0053	0.0051	0.0091	0.0053
	FNR	0.7155	0.0044	0.003	0.0031
	ACC	0.9943	0.9954	0.9958	0.9964

Conclusion

A new technique is investigated for intrusion detection system (IDS) and its performance has been evaluated on the normal and abnormal intrusion datasets. In this thesis, a new technique of NA-kNN algorithm for intrusion detection system is experimented. From the experimental results, it is seen that by using this new technique, normal and abnormal intrusion datasets could be correctly detected with 99.64% by the Manhattan distance formula. The results indicate that the data classification method has a significant impact on classification accuracy. The data used in this study was created from a limited set of programs in a single environment. The dataset can be expanded to include more variations in settings and to include more programs/processes within the Linux operating system to enable and to generalize the results for a broader set of parameters.

References

Al-Gethami, Khalid M., Mousa T. Al-Akhras, and Mohammed Alawairdhi. "Empirical evaluation of noise influence on supervised machine learning algorithms using intrusion detection datasets." *Security and Communication Networks* 2021 (2021).

Alhayali, R. A. I., Aljanabi, M., Ali, A. H., Mohammed, M. A., & Sutikno, T. (2021). Optimized machine learning algorithm for intrusion detection. *Indonesian Journal of Electrical Engineering and Computer Science*, *24*(1), 590-599.

Al-Turaiki, Isra, and Najwa Altwaijry. "A convolutional neural network for improved anomaly-based network intrusion detection." *Big Data* 9.3 (2021): 233-252.

Al-Yaseen, Wathiq Laftah, Ali Kadhum Idrees, and Faezah Hamad Almasoudy. "Wrapper feature selection method based differential evolution and extreme learning machine for intrusion detection system." *PatternRecognition* 132 (2022): 108912.

Attia, Amr, Miad Faezipour, and Abdelshakour Abuzneid. "Comparative Study of Hybrid Machine Learning Algorithms for Network Intrusion Detection." *Advances in Security, Networks, and Internet of Things*. Springer, Cham, 2021. 607-619.

Bangui, Hind, Mouzhi Ge, and Barbora Buhnova. "A hybrid machine learning model for intrusion detection in VANET." *Computing* 104.3 (2022): 503-531.

Baraneetharan, E. "Role of machine learning algorithms intrusion detection in WSNs: a survey." *Journal of Information Technology* 2.03 (2020): 161-173.

Chalé, Marc, and Nathaniel D. Bastian. "Generating realistic cyber data for training and evaluating machine learning classifiers for network intrusion detection systems." *Expert Systems with Applications* 207 (2022): 117936.

Dang, Quang-Vinh. "Improving the performance of the intrusion detection systems by the machine learning explainability". *International Journal of Web Information Systems* (2021).

Dang, Quang-Vinh. "Understanding the decision of machine learning based intrusion detection systems." *International Conference on Future Data and Security Engineering*. Springer, Cham, 2020.

Das, Anurag, Samuel A. Ajila, and Chung-Horng Lung. "A comprehensive analysis of accuracies of machine learning algorithms for network intrusion detection." *International conference on machine learning for networking*. Springer, Cham, 2020.

Dina, Ayesha S., and D. Manivannan. "Intrusion detection based on Machine Learning techniques in computer networks." *Internet of Things* 16 (2021): 100462.

Disha, Raisa Abedin, and Sajjad Waheed. "Performance analysis of machine learning models for intrusion detection system using Gini Impurity-based Weighted Random Forest (GIWRF) feature selection technique." *Cybersecurity* 5.1 (2022): 1-22.

Gamal, Merna, Hala Abbas, and Rowayda Sadek. "Hybrid approach for improving intrusion detection based on deep learning and machine learning techniques." *The International Conference on Artificial Intelligence and Computer Vision*. Springer, Cham, 2020.

Kalimuthan, C., and J. Arokia Renjit. "Review on intrusion detection using feature selection with machine learning techniques." *Materials Today: Proceedings* 33 (2020): 3794-3802.

Kumar, K. S., Nair, S. A. H., Roy, D. G., Rajalingam, B., & Kumar, R. S. (2021). Security and privacy-aware artificial intrusion detection system using federated machine learning. *Computers & Electrical Engineering*, *96*, 107440.

Megantara, Achmad Akbar, and Tohari Ahmad. "A hybrid machine learning method for increasing the performance of network intrusion detection systems." *Journal of Big Data* 8.1 (2021): 1-19.

Mendonca, Robson V., *et al*. "A lightweight intelligent intrusion detection system for industrial internet of things using deep learning algorithms." *Expert Systems* 39.5 (2022): e12917.

Meryem, Amar, and Bouabid EL Ouahidi. "Hybrid intrusion detection system using machine learning." *Network Security* 2020.5 (2020): 8-19.

Nazir, Anjum, and Rizwan Ahmed Khan. "A novel combinatorial optimization based feature selection method for network intrusion detection." *Computers & Security* 102 (2021): 102164.

Rajesh, M. V. "Intensive analysis of intrusion detection methodology over Mobile Adhoc Network using machine learning strategies." *Materials Today: Proceedings* 51 (2022): 156-160.

Salih, Azar Abid, and Adnan Mohsin Abdulazeez. "Evaluation of classification algorithms for intrusion detection system: A review." *Journal of Soft Computing and Data Mining* 2.1 (2021): 31-40.

Seniaray, Sumedha, and Rajni Jindal. "Machine Learning-Based Network Intrusion Detection System." *Computer Networks and Inventive Communication Technologies*. Springer, Singapore, 2022. 175-187.

Sharma, Mrinal, and C. R. S. Kumar. "Machine learning-based smart surveillance and intrusion detection system for national geographic borders." *Artificial Intelligence and Technologies*. Springer, Singapore, 2022. 165-176.

Singh, Geeta, and Neelu Khare. "A survey of intrusion detection from the perspective of intrusion datasets and machine learning techniques." *International Journal of Computers and Applications* (2021): 1-11.

Susilo, Bambang, and Riri Fitri Sari. "Intrusion detection in IoT networks using deep learning algorithm." *Information* 11.5 (2020): 279.

Varanasi, VenkataRamani, and Shaik Razia. "Network Intrusion Detection using Machine Learning, Deep Learning-A Review." *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 2022.

Verkerken, M., D'hooge, L., Wauters, T., Volckaert, B., & De Turck, F. (2022). Towards model generalization for intrusion detection: Unsupervised machine learning techniques. *Journal of Network and Systems Management*, *30*(1), 1-25.