# A Lightweight Hybrid Intrusion Detection Framework using Machine Learning for Edge-Based IIoT Security

Azidine Guezzaz
Technology Higher School Essaouira, Cadi Ayyad University, Morocco
a.guzzaz@gmail.com

Mourade Azrour
IDMS team, Faculty of Sciences and Technics, Moulay Ismail University of Meknès, Morocco
mo.azrour@umi.ac.ma

Said Benkirane
Technology Higher School Essaouira, Cadi Ayyad University, Morocco
sabenkirane@gmail.com

Mouaad Mohy-Eddine
Technology Higher School Essaouira, Cadi Ayyad University, Morocco
mouaadmohyeddine@gmail.com

Hanaa Attou
Technology Higher School Essaouira, Cadi Ayyad University, Morocco
attouhanaa39@gmail.com

Maryam Douiba
Technology Higher School Essaouira, Cadi Ayyad University, Morocco
maryam.douiba5@gmail.com

**Abstract:** *Due to the development of cloud computing and Internet of Things (IoT) environments, such as healthcare systems, telecommunications and Industry 4.0 or Industrial IoT (IIoT) many daily services are transformed. Therefore, Security issues become useful to better protect these novel technologies. IIoT security represents a real challenge for industry actors and academic research. A set of security approaches, such as intrusion detection are integrated to improve IIoT environments security. Hence, an Intrusion Detection System (IDS) aims to monitor, detect an intrusion in real time and then make reliable decisions. Many recent IDS incorporate Machine Learning (ML) techniques to improve their Accuracy (ACC), precision and Detection Rate (DR). This paper presents a hybrid IDS for Edge-Based IIoT Security using ML techniques. This new hybrid framework is based on misuse and anomaly detection using K-Nearest Neighbor (K-NN) and Principal Component Analysis (PCA) techniques. Specifically, the K-NN classifier has been incorporated to improve detection accuracy and make effective decision and the PCA is used for an enhanced feature engineering and training process. The obtained results have proven that our proposed Framework presents many advantages compared with other recent models. It gives good results with 99.10% ACC, 98.4% DR 2.7% False Alarm Rate (FAR) on NSL-KDD dataset and 98.2% ACC, 97.6% DR, 2.9% FAR on Bot-IoT dataset.*

**Keywords:** *IoT security, edge-based IIoT, intrusion detection, ML, K-NN, PCA. NSL-K, Bot-IoT.*

*Received August 6, 2021, accepted December 9, 2021*
*https://doi.org/10.34028/iajit/19/5/14*

## 1. Introduction

Due to the rapid growth of Internet of Things (IoT)-loud and Industry 4.0 technologies, the security and privacy issues are becoming roughly useful to protect networks and data. The main goal of these new technologies is to facilitate and deliver accurately various services for companies and persons [6, 8, 9, 10, 16, 19, 43]. The industry 4.0 also called Industrial IoT (IIoT) is a recent technology allowing the use of smart sensors and actuators to enhance industrial processes. The data is exchanged between these intelligent edge devices using various communications infrastructure [53]. Due to the increase of IIoT environments, the security needs represent one of their actual challenges. The IIoT security integrates solutions to control and protect accurately data and several edge devices by providing sophisticated approaches [11, 33, 43, 45]. Intrusion Detection Systems (IDSs) are among recent technologies implemented to detect intrusions and undesirable activities using misuse or anomaly detection [21, 23, 26, 34, 36, 39]. The hybrid approach is obtained by combining advantages of both methods. It aims to increase detection rate and accuracy of IDS [18, 26, 27]. We distinguish also between Network-based Intrusion Detection System (NIDS) and Host-based Intrusion Detection System (HIDS) [26, 28]. IDS suffer from much of limits, such as real-time detection, big volume of data, and generated alarm that can decrease IDS accuracy and its detection rate [29]. Intrusion detection approaches incorporated in Industry 4.0 environments are classified into knowledge-based and anomaly-based. Knowledge-based methods cannot detect unknown intrusive traffic [53].

Various Machine Learning (ML) techniques are used to allow learning from training data to obtain and build reliable models. The new instances will be detected correctly based on models obtained during the training process.

Hence, the generalized learning provides the possibility to process data that have never been

examined before. In fact, standard ML methods are more suitable for an IIoT edge device due to the learning process that is relatively short and the computational speed of the system which is extremely limited. The intrusion detection represents an ongoing research field which aims to design and validate enhanced approaches by improving feature engineering and data quality [3, 7, 22, 25, 30, 38, 44], but also to enhance classifiers [2, 15, 27, 31] for making good decisions.

In this work, we design and validate a robust hybrid intrusion detection framework, called Principal Component Analysis (PCA) and K-Nearest Neighbor (K-NN) IDS (PK-IDS). Specifically, the PCA method that applied on collected traffic to reduce features and enhancing features engineering. The K-NN algorithm is incorporated to build a binary classifier model to make reliable decisions and reduce false alarm. The main goal of this research is to validate two fundamental contributions. First, we implement a feature engineering solution based on PCA algorithm aiming to enhance data preparation process; we also validate a K-NN classifier model to achieve accurate intrusion detection approach. Second, we proposed a robust hybrid framework using two detection methods: misuse detection which achieved by integrating Snort intrusion detection system [17, 26, 45] and anomaly detection which based on our proposed PCA feature engineering and K-NN classifier. Experimental results on Bot-IoT dataset and NSL-KDD dataset demonstrates that our proposed approach presents good performances in term of Accuracy (ACC) and Detection Rate (DR). Thereby, the PK-IDS framework gives many advantages when compared to other detection approaches according to many measures including accuracy and data quality.

The rest of the article is organized as follows. In section 2, related works on intrusion detection in IoT and IIoT environments are cited, specifically those incorporated ML techniques to enhance detection approaches. The section 3 describes various suggested solutions of the PK-IDS framework. In section 4, we discuss results that obtained in experimental setting, but also a comparison of the new proposed framework and others. Finally, the paper is achieved with a conclusion and perspectives.

## 2. Related Works

In this section, we present certain research works of intrusion detection; especially those introduce ML techniques to secure data and networks.

Detecting intrusions in such an IoT environment is a critical and difficult security issue. Various IDSs integrating ML and Deep Learning (DL) have been designed for different types of attacks. Hence, a range of contributions have proposed the enhanced intrusion detection approaches using ML and DL techniques. Specifically, DL methods [1, 4, 5, 7, 12, 16, 19, 22, 24, 32, 40], Multi-Layer Perceptron (MLP) [25], K-NN [50], Support Vector Machine (SVM) [1, 25, 37], Naïve Bayes (NB) [1, 25, 42, 44, 48] Decision Tree (DT) [1, 30, 35, 41], Random Forest (RF) [30], fuzzy clustering [5], and Reinforcement based Learning (RL) [46]. Ahmim *et al*. [2] proposed a hybrid IDS model which combining classifier model based on decision tree, REP tree, JRIP algorithm and forest PA. The performances of a novel model are evaluated using CICIDS2017 dataset. On the other side, a set of feature engineering techniques are proposed to obtain efficient IDSs by improving preprocessing and data quality aiming to train and build accurate classifiers [7, 25]. In 2020, Mighan and Kahani [40] proposed a novel intrusion detection model using DL method and also integrate SVM, RF, DT and NB techniques to make scalable IDS. To evaluate the model, UNB ISCX 2012 dataset is used. Hassan *et al*. [32] develops a hybrid DL model for efficient intrusion detection in big data environment that integrates convolutional neural network and long short-term memory network methods. The performances are validated using ISCX2012 and UNSW-NB15 datasets. Sethi *et al*. [46] proposed a robust model of IDS using reinforcement learning based approach. The NSL-KDD dataset, UNSW-NB15 dataset and AWID dataset are used to validate performances. Elmasry *et al*. [20] develops a DL model for network intrusion detection using a double PSO metaheuristic. This new model is evaluated on CICIDS2017 and NSL-KDD datasets and achieves good results. Chiba *et al*. [16] proposed intelligent IDS based on DL method to secure cloud environment. It combines several of ML algorithms. Hence, a genetic algorithm is implemented to enhance feature engineering process and DL for classification intrusions. KDD CUP 1999 dataset is used to validate the model performances. Guezzaz *et al*. [27] developed a recent IDS approach integrating MLP classifier. The approach includes PcapSocks sniffer to collect traffic from networks and MLP classifier to classify data into intrusion or normal activity.

Today most of proposed IDS focuses on feature selection or reduction [2, 38] aiming to avoid and eliminate irrelevant features that can decreases IDS performances. Furthermore, the main goal of feature engineering operations is to identify and select useful reduced inputs to build effective IDS [7, 44]. Gu *et al*. [25] proposed a novel IDS model which implements SVM ensemble for classification and NB algorithm for feature augmentation. UNSW-B15, NSL-KDD, CICIDS2017 and Kyoto2006+ datasets are used to validate performances of this recent approach. Abd *et al*. [1] developed an intrusion detection model which integrated NB and DL technique. The new IDS implements genetic algorithm to select useful features. Ayo *et al*. [7] proposed network intrusion detection model using DL method optimized with rule-based

hybrid feature selection. The UNSW-NB15 dataset is used validate the model. Prasad *et al*. [44] proposed new IDS which work on subset of features by extracting significant features using the probabilistic method. The BRS method is implemented to categorize samples into normal, intermediary, and abnormal category based on rough set. The model is trained and tested on NLS-KDD, UNSW-NB15 and CICIDS2017 datasets. Kumar *et al*. [38] proposed an enhanced IDS using multi-class SVM to build a reliable classifier and multi-Linear dimensionality reduction to improve data quality before making decision. The NSL-KDD dataset is used to validate system. However, some recent works focus effectively to integrate ML and DL techniques for intrusion detection in IoT environments. Chaabouni *et al*. [14] proposed a network based IDS in IoT environment based on ML techniques with the benchmark datasets and open source intrusion detection tools are discussed. Wu *et al*. [52] developed a hybrid IDS model combining a stacked auto-encoder with SVM and Kernel approximation in IoT. The NSL-KDD dataset is used to evaluate the performances of this model. Bagaa *et al*. [13] proposed a novel ML model for securing aspects according to IoT systems. This anomaly detection model is based on SVM classifier with ACC 99.71% and DR 98.8% on NSL-KDD dataset. Verma and Ranga [51] presented an anomaly detection model using ML algorithms to secure IoT environments against Denial-of-Service (DoS attacks. CIDDS-001, UNSW-NB15 and NSL-KDD datasets are used to evaluate model. From results performances and comparison: CART shows results of ACC 96.74% and AB presents DR 97.5%. For sensitivity, RF and XGB present 97.3% metrics. In 2020, Sarker *et al*. [45] proposed a ML intrusion detection model based Cyber security, called IntruDTree. The model takes into account the important security by implementing feature selection method. Compared with previous works, the IntruDTree gives good results: 98% ACC and DR 98%. Kirana *et al*. [37] build ML classifiers to identify Main-In-The-Middle attacks in IoT networks. The proposed model includes NB, SVM, and Adabost to classify data into normal and attack. Experiments evaluations are performed on captured dataset sensor 480. The performance results give ACC 98% and 98% DR with SVM, ACC 97% and DR 96% with NB, ACC 98% and 98% DR with Adabost. To design and validate a robust intrusion detection approach for Edge-Based IIoT Security is still a difficult issue where many considerations have to be taken into account. From overview and related works, it is proven that the learning methods and feature engineering are two important tasks which determine the effectiveness and accuracy of IDS.

## 3. Our Proposed Approach

This section includes a series of proposed solutions to validate our PK-IDS system through the implementation of powerful technical features and classification mechanisms.

### 3.1. Proposed Scheme

According to the standard IDS constituents revealed in [24, 27, 36], our hybridized framework is designed in four parts: data gathering, the pre-processing, the decision making and the response. Accordingly, our suggested system, which is depicted in Figure 1, represents a hybrid model which is based on signature detection by using the Snort IDS and anomaly detection by using the K-NN classifier.

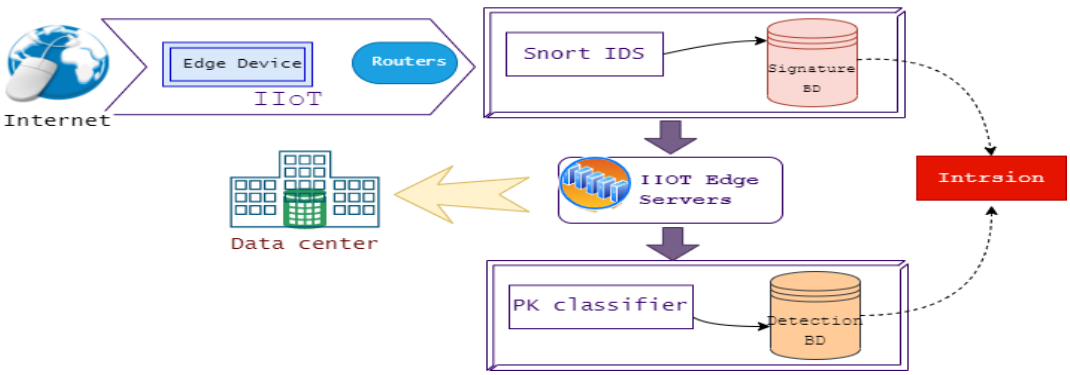



Figure 1. Proposed scheme of IDS for edge-based IIoT security.

The newly developed architecture represents a centralized architecture that will be deployed on a central device for the IIoT environment, for instance a central server equipped with a very high powerful memory and a high performance computational processor. The new PK-IDS model is a multi-functional hybrid network intrusion detection tool that–combines Snort IDS-based malicious behavior detection and anomaly detection using PCA technology and the K-NN classifier. Figure 2 shows the implementation steps of our framework in detail.

The proposed approach represents a centralized

architecture that will be installed on a central device for the IIoT environment, such as a central server characterized by a very powerful memory and an efficient calculation rate which allows our IDS framework to perform efficient processing, especially since it integrates complex ML techniques which require very high time consumption and efficient storage. However, a pre-processing and feature engineering process have been performed on data collection to increase time complexity and decrease edge-based IIoT resources utilization.

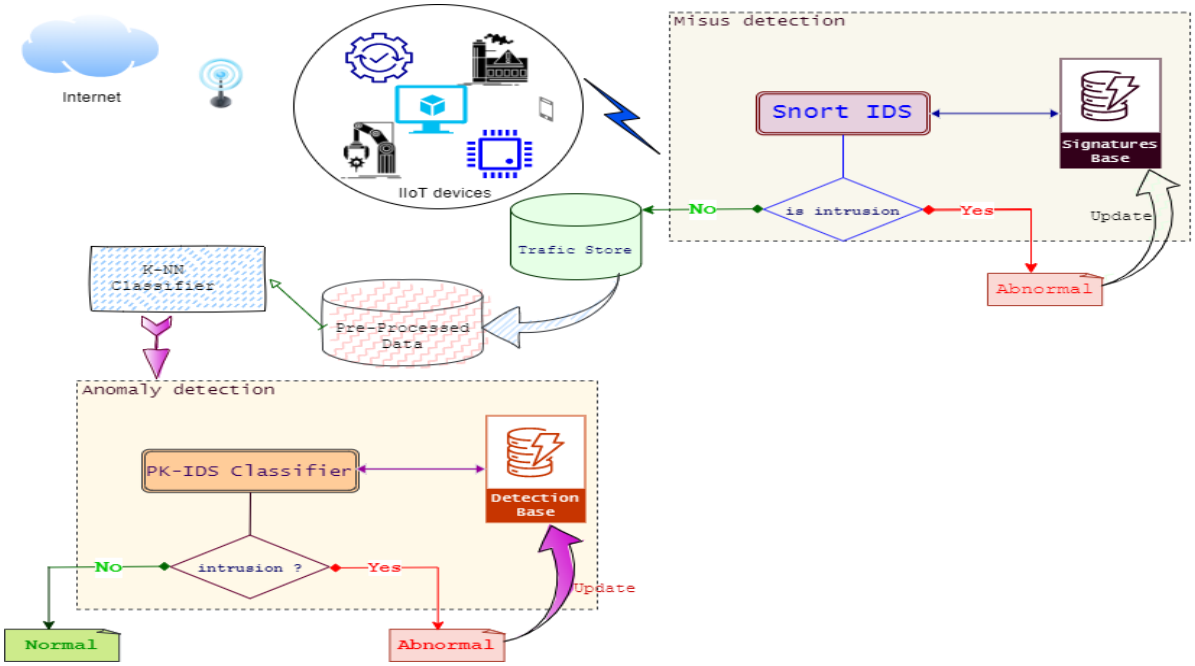



Figure 2. Novel PK-IDS framework for edge-Based IIoT security.

### 3.2. Detailed Description of Solutions

As shown in Figure 2, the proposed architecture for the PK-IDS framework consists of five actions: abuse detection process, preprocessing and normalization, feature engineering, training and validation, and classification.

1. Misuse detection process: this procedure is achieved by Snort IDS [17, 47], once the packet is gathered; it is analyzed and compared with signature base which contains predefined network attack rules. Snort matches the collected packet with rules. Once the attack is detected, an alert is sent to inform the nature of attack, the signature base is updated and a convenient decision is made. The normal packets are not detected and forwarded to next procedures of framework.
2. Pre-processing and normalization: our proposed approach focuses on pre-processing and normalization to improve data quality. The pre-processing is a useful task aiming to eliminate noise and clean data. Also, the data are normalized in range [0, 1] using min max method to avoid undesirable influence of high weight of values features. The Equation (1) is used to find the new value. Hence, we make the values of each feature run from 0 to 1. If the lowest value of a given feature x is min and the highest value is max, we convert each value of x to:

$$(value(x) - \min)/(max - \min) \quad (1)$$

3. Feature engineering: in the implementation step, we propose to extract samples of dataset to avoid some drawbacks such as, processing and big volume of data. There are many strategies in which the number of features can be reduced before a dataset is used for training and validation of classifier model. In this case, we use PCA for achieving this task. The PCA is a statistical technique aiming to obtain lower dimensional form of the original data by reducing high and complex dimensionality without losing any significant information. The reduction process decrease training time and computation cost without affecting the performances of results. The feature engineering process selects useful features of original data and enhances data quality. This allows us to build an accurate classifier for our PK-IDS framework.
4. Training and validation: to validate our proposed model, we implement the 10-fold cross validation method recommended in [23]. The procedure aims to split randomly full dataset into ten parts with same size. Nine parts are used to train model and the last part in the test. The performances are carried out by repeating this procedure ten times used to train and build an accurate classifier able to discover intrusions within traffic network. Only the features meeting a specified criterion are used in training and validation of the model.

5. Classification: the classification is the last process, which assigns a class to an instance. Therefore, the obtained classifier from the training and validation step allows predicting classes of new instances. For this, we use the K-NN algorithm.

## 4. Experimental Setting

In this section, we present the background of the experiments and the datasets employed in the actual experimental study. We also address the performance of the obtained results and the comparison of the proposed method with previously published works.

### 4.1. Dataset and Environment

The evaluation of datasets has a main role in validation of intrusion detection approaches. Many appropriate public datasets are available to evaluate any IDS based on ML techniques [36, 41, 49]. In this paper, the Bot-IoT dataset and NSL-KDD dataset are used to train, evaluate and validate our proposed model. They are a public and most used datasets, especially for IoT network traffic aiming to produce reliable evaluation results of research works.

The Bot-IoT dataset has been created in the Cyber Range Lab of The center of UNSW Canberra Cyber by setting up a network environment which combines normal and botnet traffic. It contains DoS, DDoS and service scan including different files, such as Pcap files which have 69.3 GB in size with 72.000.000 records and CSV files with 16.7 GB in size. Those files are separated according to category and subcategory [36].

NSL-KDD dataset was created from KDD cup 99 dataset [40, 49]. It includes 125,973 records used in training step and 22,544 for test step. It contains 22 training instances attacks and 41 features which 21 of them describe connection itself and 19 for nature of connection of the same host. The novelty and instances volume of NSL-KDD dataset make it very practical. The Bot-IoT dataset and NSL-KDD dataset used in this research work are available at https://web.archive.org/web/20150205070216/http://nsl.cs.unb.ca/NSL-KDD/ and https://research.unsw.edu.au/projects/bot-iot-dataset respectively. The experimental study and evaluation of our work were all performed on a computer running windows 7 professional 64 bits with a Core-i7 2700K CPU@ 2.50 GHz and 32 GB of DDR3. The python version 3.8.0 is used to implement and validate PCA features engineering and K-NN model training.

### 4.2. Discussion of Results

After pre-processing and NB features engineering process on collected traffic, the performances evaluation of K-NN classifier model are performed using ACC, DR rate, FAR and f-score measures. For this, used confusion matrix is described in Table 1.

Table 1. Confusion matrix.

| Actual class | Predicted class | |
|---|---|---|
| | Attack | Normal |
| Attack | TP | FN |
| Normal | FP | TN |

The evaluation metrics of classifiers considered in this study are calculated as follow:

- ACC is the ratio of instances that are correctly predicted as normal or attack to the overall number of instances in test set. It is obtained from Equation (2).

$$\text{ACC} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (2)$$

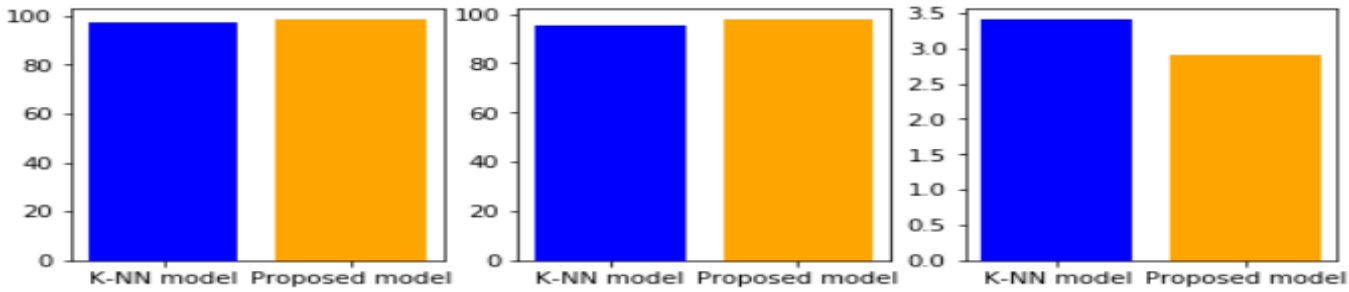
- DR indicates the ratio of the number of instances that are correctly classified as attack to the total number of attack instances present in test set. It is calculated using Equation (3).

$$\text{DR} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

- FAR represents the ratio of instances which is categorized as attack to the overall number of instances of normal behavior. It is obtained from Equation (4).

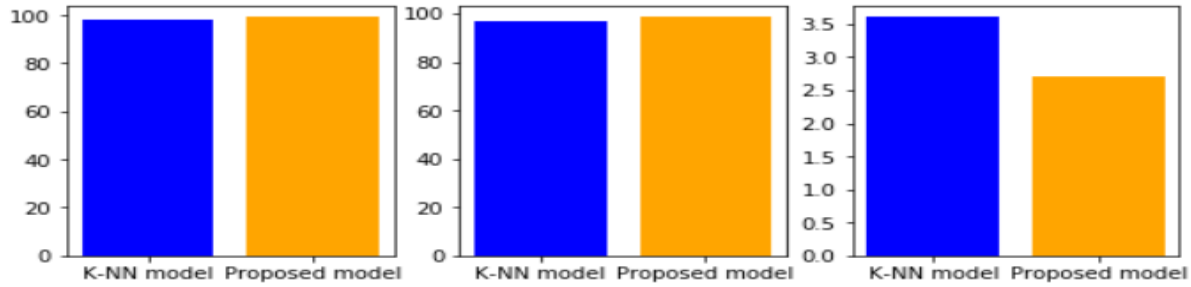
$$\text{FAR} = \frac{\text{FP}}{\text{FP+TN}} \quad (4)$$

In this work, we start with comparing detection evaluation of our proposed approach and K-NN model only. The results obtained in Figures 3 and 4 show this comparison according to ACC, DR and FAR on Bot-IoT dataset and NSL-KDD dataset.




a) ACC (%). b) DR (%). c) FAR (%).

Figure 3. Performances metrics of K-NN model and our proposed model on Bot-IoT dataset.




a) ACC (%). b) DR (%). c) FAR (%).

Figure 4. Performances metrics of K-NN model and our proposed model on NSL-KDD dataset.

Tables 2, and 3 show the performance comparison of various anomaly based IDS models on the original binary class NSL-KDD dataset and the dimensionally reduced NSL-KDD dataset obtained after applying PCA with 98% variance retention, respectively. It can be observed from these tables that the ACC and DR of the anomaly based IDSs obtained using the dimensionally reduced dataset is comparable to that obtained using the original higher dimensional dataset.

The Figure 3-a), and Figure 4-a) show that ACC of our novel approach is specifically better than model based on K-NN only, Figure 3-b) and Figure 4-b) show the DR of both IDS. It validates that the DR of proposed IDS model is higher than the IDS based on K-NN only on Bot-IoT dataset and NSL-KDD dataset.

The results demonstrated above are summarized in Tables 3, and 4. They show that our proposed model can reach significant performances than k-NN only. For Bot-IoT dataset, the ACC of our proposed model achieves 98.2% while K-NN only exceeds 97%. In terms of DR and FAR, our proposed model obtains 97.6% and 2.9% respectively, while decision tree only presents DR 95.5% and FAR 3.4%. For NSL-KDD dataset, our proposed model indicates high performances in terms of ACC 99.10%, DR 98.4% and FAR 2.7%. Besides, the K-NN only gives ACC 98%, DR 96.7% and FAR 3.6%.

Table 2. Performances metrics of K-NN and proposed model on Bot-IoT dataset.

| | ACC (%) | DR (%) | FAR (%) |
|---|---|---|---|
| **K-NN** | 97% | 95.5% | 3.4% |
| **Proposed model** | 98.2% | 97.6% | 2.9% |

Table 3. Performances metrics of K-NN and proposed model on NSL-KDD dataset.

| | ACC (%) | DR (%) | FAR (%) |
|---|---|---|---|
| **K-NN** | 98% | 96.7 | 3.6% |
| **Proposed model** | 99.10 | 98.4 | 2.7% |

The results obtained validate that our approach gives great detection capability in terms of ACC, DR and FAR. Specifically, they demonstrate that the performances metrics of our proposed model are higher on NSL-KDD dataset but low on Bot-IoT dataset. According to the evaluation performances; our proposed IDS model can reach great performances. The comparison with model which uses K-NN only indicates the effectiveness of our network intrusion detection approach. Concretely, our proposed intrusion detection model is specified by high performances of ACC, DR, and FAR. Furthermore, we perform a comparison between our IDS and other recent intrusion detection approaches based on on Bot-IoT dataset and on NSL-KDD dataset. Typically, the recent works that integrate ML techniques such as: SVM, NB, Adabost and CART. The comparison results are presented in Table 4.

Table 4. Comparison with previous works on NSL-KDD.

| | Method | ACC (%) | DR (%) |
|---|---|---|---|
| **Bagaa *et al.* [13]** | | 99.71 | 98.8 |
| **Verma and Ranga [51]** | CART | 96.74 | - |
| | AB | - | 97.5 |
| **Sarker *et al.* [45]** | | 98 | 98 |
| **Mukherjee *et al.* [42]** | SVM | 98 | 98 |
| | NB | 97 | 96 |
| | Adabost | 98 | 98 |
| **Proposed model** | PCA, K-NN | 99.10 | 98.4 |

From the obtained results, we conclude that our proposed IDS approach is relevant, achieves important performances and gives relevant training by implementing fast data quality techniques. Using Bot-IoT dataset and NSL-KDD dataset, it is proven that our approach is reliable and reaches good results compared with other models. The novel approach can be integrated and used to secure various environments such as IoT environment and cloud computing.

## 5. Conclusions and Perspectives

This paper has presented an effective hybrid intrusion detection framework using K-NN classifier and PCA technique for feature engineering. The new hybrid framework called PK-IDS is a hybrid approach that integrates Snort IDS for misuse detection and K-NN classifier to improve anomaly detection approach. Preparation of data is a process that is handled according to non-standardization and heterogeneity of data. We have performed PCA technique for feature engineering to increase training step, accuracy and detection rate of built model. The PK-IDS framework is validated by implementing various proposed solutions. The empirical study of our novel hybrid approach is evaluated using Bot-IoT and NSL-KDD datasets. Compared with other approaches, the novel framework gives robust performances and accuracy. In fact, the

design of IDS for edge-based IIoT security is a difficult task and faces many problems, such as how to place IDS. Our future work will empower PK-IDS framework by incorporating some sophisticated techniques of artificial intelligence taking into consideration various edge-based IIoT characteristics.

## References


[1] Abd S., Alsajri M., and Ibraheem H., "Rao-SVM Machine Learning Algorithm for Intrusion Detection System," *Iraqi Journal for Computer Science and Mathematics*, vol. 1, no. 1, pp. 23-27, 2020.

[2] Ahmim A., Maglaras L., Ferrag M., Derdour M., and Janicke H., "A Novel Hierarchical Intrusion Detection System Based on Decision Tree and Rules-Based Models," *in Proceeding of the International Conference on Distributed Computing in Sensor Systems*, Santorini, pp. 228-233, 2019.

[3] Alazzam H., Sharieh A., and Sabri K., "A Feature Selection Algorithm for Intrusion Detection System Based on Pigeon Inspired Optimizer," *Expert Systems with Applications*, vol.148, pp. 113249, 2020.

[4] Aldweesh A., Derhab A., and Emam A., "Deep Learning Approaches for Anomaly-Based Intrusion Detection Systems: A Survey, Taxonomy, and Open Issues," *Knowledge-Based Systems*, vol. 189, pp. 105124, 2020.

[5] Amini M., Rezaeenour J., and Hadavandi E., "A Neural Network Ensemble Classifier for Effective Intrusion Detection Using Fuzzy Clustering and Radial Basis Function Networks," *International Journal on Artificial Intelligence Tools*, vol. 25, no. 02, pp. 1550033, 2016.

[6] Atzori L., Iera A., and Morabito G., "The Internet of Things: A Survey," *Computer Networks*, vol. 54, no. 15, pp. 2787-2805, 2010.

[7] Ayo F., Folorunso S., Abayomi-Alli A., Adekunle A., and Awotunde J., "Network Intrusion Detection Based on Deep Learning Model Optimized with Rule-based Hybrid Feature Selection," *Information Security Journal: A Global Perspective*, vol. 29, no. 6, pp. 267-283, 2020.

[8] Azrour M., Mabrouki J., Farhaoui Y., and Guezzaz A., "Experimental Evaluation of Proposed Algorithm for Identifying Abnormal Messages in SIP Network," *in Intelligent Systems in Big Data, Semantic Web and Machine* Learning, pp. 1–10, 2021.

[9] Azrour M., Mabrouki J., and Chaganti R., "New Efficient and Secured Authentication Protocol for Remote Healthcare Systems in Cloud-IoT," *Security and Communication Networks*, 2021.

[10] Azrour M., Mabrouki J., Guezzaz A., and Farhaoui Y., "New Enhanced Authentication Protocol for Internet of Things," *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 1-9 2021.

[11] Azrour M., Mabrouki J., Guezzaz A., and Kanwal A., "Internet of Things Security: Challenges and Key Issues," *Security and Communication Networks*, vol. 2021, 2021.

[12] Azrour M., Mabrouki J., Fattah G., Guezzaz A., and Aziz F., "Machine Learning Algorithms for Efficient Water Quality Prediction Model," *Modeling Earth Systems and Environment*, vol. 8, no. 2, pp. 27930-2801, 2021.

[13] Bagaa M., Taleb T., Bernabe J., and Skarmeta A., "A Machine Learning Security Framework for Iot Systems," *IEEE Access*, vol. 8, pp. 114066-114077, 2020.

[14] Chaabouni N., Mosbah M., Zemmari A., Sauvignac C., and Faruki P., "Network Intrusion Detection for IoT Security based on Learning Techniques," *IEEE Communications Surveys and Tutorials*, vol. 31, no. 3, pp. 2671-2701, 2019.

[15] Chanal P. and Kakkasageri M., "Security and Privacy in IoT: A Survey," *Wireless Personal Communications*, vol. 115, no. 2, pp. 1667-1693, 2020.

[16] Chiba Z., Abghour N., Moussaid K., Elomri A., and Rida M., "Intelligent Approach to Build A Deep Neural Network Based IDS for Cloud Environment Using Combination of Machine Learning Algorithms," *Computers and Security*, vol. 86, pp. 291-317, 2019.

[17] Chiba Z., Abghour N., Moussaid K., and Rida M., "A Cooperative and Hybrid Network Intrusion Detection Framework in Cloud Computing Based on Snort and Optimized Back Propagation Neural Network," *Procedia Computer Science*, vol. 83, pp. 1200-1206, 2016.

[18] Çavuşoğlu Ü., "A New Hybrid Approach for Intrusion Detection Using Machine Learning Methods," *Applied Intelligence*, vol. 49, no. 7, pp. 2735-2761, 2019.

[19] Colakovic A. and Hadÿzialic M., "Internet of Things (IoT): A Review of Enabling Technologies, Challenges, and Open Research Issues," *Computer Networks*, vol. 144, pp. 17-39, 2018.

[20] Elmasry W., Akbulut A., and Zaim A., "Evolving Deep Learning Architectures for Network Intrusion Detection Using A Double PSO Metaheuristic," *Computer Networks*, vol. 168, pp. 107042, 2020.

[21] Fang W., Tan X., and Wilbur D., "Application of Intrusion Detection Technology in Network Safety Based on Machine Learning," *Safety Science*, vol. 124, pp. 104604, 2020.

[22] Farhan B. and Jasim A., "A Survey of Intrusion Detection Using Deep Learning in Internet of Things," *Iraqi Journal for Computer Science and*

*Mathematics*, vol. 3, no. 1, pp. 83-93, 2022.

[23] Fernandes G., Rodrigues J., and Carvalho L., "A Comprehensive Survey on Network Anomaly Detection," *Telecommunication Systems*, vol. 70, no. 3, pp. 447-489, 2019.

[24] Ferrag M., Maglaras L., Moschoyiannis S., and Janicke H., "Deep Learning for Cyber Security Intrusion Detection: Approaches, Datasets, and Comparative Study," *Journal of Information Security and Applications*, vol. 50, pp. 102419, 2020.

[25] Gu J., Wang L., Wang H., and Wang S., "A Novel Approach to Intrusion Detection Using SVM Ensemble with Feature Augmentation," *Computers and Security*, vol. 86, pp. 53-62, 2019.

[26] Guezzaz A., Asimi A., Asimi Y., Tbatou Z., and Sadqi Y., "A Global Intrusion Detection System using PcapSockS Sniffer and Multilayer Perceptron Classifier," *International Journal of Network Security*, vol. 21, no. 3, pp. 438-450, 2019.

[27] Guezzaz A., Asimi Y., Azrour M., and Asimi A., "Mathematical Validation of Proposed Machine Learning Classifier for Heterogeneous Traffic and Anomaly Detection," *Big Data Mining and Analytics*, vol. 4, no. 1, pp. 18-24, 2021.

[28] Guezzaz A., Asimi A., Asimi Y., Azrour M., and Benkirane S., "A distributed intrusion detection approach based on machine learning techniques for a cloud security," *in Intelligent Systems in Big Data, Semantic Web and Machine Learning, Springer*, pp. 85-94, 2021.

[29] Guezzaz A., Asimi A., Sadqi Y., Asimi Y., and Tbatou Z., "A New Hybrid Network Sniffer Model Based on Pcap Language and Sockets (Pcapsocks)," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 2, 2016.

[30] Guezzaz A., Benkirane S., Azrour M., and Khurram S., "A Reliable Network Intrusion Detection Approach Using Decision Tree with Enhanced Data Quality," *Security and Communication Networks*, vol. 2021, 2021.

[31] Hadi A.,"Performance Analysis of Big Data Intrusion Detection System over Random Forest Algorithm," *International Journal of Applied Engineering Research*, vol. 13, no. 2, pp. 1520-1527, 2018.

[32] Hassan M., Gumaei A., Alsanad A., Alrubaian M., and Fortino G., "A Hybrid Deep Learning Model for Efficient Intrusion Detection in Big Data Environment," *Information Sciences*, vol. 513, pp. 386-396, 2020.

[33] Ingham M., Marchang J., and Bhowmik D., "IoT Security Vulnerabilities and Predictive Signal Jamming Attack Analysis in LoRaWAN," *IET Information Security*, vol. 14, no. 4, pp. 368-379, 2020.

[34] Ji S., Jeong B., Choi S., and Jeong D., "A Multi-Level Intrusion Detection Method for Abnormal Network Behaviors," *Journal of Network and Computer Applications*, vol. 62, pp. 9-17, 2016.

[35] Kevric J., Jukic S., and Subasi A., "An Effective Combining Classifier Approach Using Tree Algorithms for Network Intrusion Detection," *Neural Comput and Applic*, vol. 28, no. 1, pp. 1051-1058, 2017.

[36] Khraisat A., Gondal I., Vamplew P., and Kamruzzaman J., "Survey of Intrusion Detection Systems: Techniques, Datasets and Challenges," *Cybersecurity*, vol. 2, no. 1, pp. 1-22, 2019.

[37] Kirana K., Devisettya R., Kalyana N., Mukundini K., and Karthi R., "Building a Intrusion Detection System for IoT Environment using Machine Learning Techniques," *Procedia Computer Science*, vol. 171, pp. 2372-2379, 2020.

[38] Kumar B., Raju M., Vardhan B., "Enhancing the Performance of an Intrusion Detection System through Multi-Linear Dimensionality Reduction and Multi-Class SVM," *International Journal of Intelligent Engineering and Systems*, vol. 11, no. 1, pp. 181-192, 2018.

[39] Masdari M. and Khezri H., "A Survey and Taxonomy of the Fuzzy Signature-based Intrusion Detection Systems," *Applied Soft Computing*, vol. 92, pp. 106301, 2020.

[40] Mighan S. and Kahani M., "A Novel Scalable Intrusion Detection System Based on Deep Learning," *International Journal of Information Security*, vol. 20, no. 3, pp. 387-403, 2020.

[41] Moustafa N. and Slay J., "The Evaluation of Network Anomaly Detection Systems: Statistical Analysis of the Unsw-Nb15 Data Set and the Comparison with the Kdd99 Data Set," *Information Security Journal: A Global Perspective*, vol. 25, no. 1-3, pp. 18-31, 2016.

[42] Mukherjee S., Sharma N., "Intrusion Detection using Naive Bayes Classifier with Feature Reduction," *Procedia Technology*, vol. 4, pp. 119-128. 2012.

[43] Noor M. and Hassan W., "Current Research on Internet of Things (IoT) Security: A Survey," *Computer Networks*, vol. 148, pp. 283-294, 2018.

[44] Prasad M., Tripathi S., and Dahal K., "An Efficient Feature Selection Based Bayesian and Rough Set Approach for Intrusion Detection," *Applied Soft Computing*, vol. 87, pp. 105980, 2020.

[45] Sarker I., Abushark Y., Alsolami F., and Khan A., "IntruDTree: A Machine Learning Based Cyber Security Intrusion Detection Model," *Symmestry*, vol. 15, no. 5, pp. 754, 2020.

[46] Sethi K., Rupesh E., Kumar R., Bera P., and Madhav Y., "A Context-Aware Robust Intrusion Detection System: A Reinforcement Learning-Based Approach," *International Journal of*

*Information Security*, vol. 19, no. 6, pp. 657-678, 2020.
[47] Snort-Network Intrusion Detection and Prevention System., https://www.snort.org/ Last Visited, 2022.
[48] Tabash M., Abd-Allah M., and Tawfik B., "Intrusion Detection Model Using Naive Bayes and Deep Learning Technique," *The International Arab Journal of Information Technology*, vol. 17, no. 2, pp. 215-224, 2020.
[49] Tavallaee M., Bagheri E., Lu W., and Ghorbani A., "A Detailed Analysis of the KDD CUP 99 Dataset," *in Proceedings of the IEEE Symposium on Computational Intelligence for Security and Defense Applications*, Ottawa, pp.1-6, 2009.
[50] Topirceanua A. and Grosseckb G., "Decision Tree Learning Used for the Classification of Student Archetypes in Online Courses," *Procedia Computer Science*, vol. 112, pp. 51-60, 2017.
[51] Verma A. and Ranga V., "Machine Learning Based Intrusion Detection Systems for IoT Applications," *Wireless Personal Communications*, vol. 111, no. 4, pp. 2287-2310, 2019.
[52] Wu Y., Lee W., Gong X., and Wang H., "A Hybrid Intrusion Detection Model Combining SAE with Kernel Approximation in Internet of Things," *Sensors*, vol. 20, no. 19, pp. 5710, 2020.
[53] Yao H., Gao P., Zhang P., Wang J., Jiang C., and Lu L.," Hybrid Intrusion Detection System for Edge-Based IIoT Relying on Machine-Learning Aided Detection" *IEEE Network*, vol. 33, no. 5, pp. 75-81, 2018.

**Azidine Guezzaz** received his Ph.D from Ibn Zohr University Agadir, Morocco in 2018. He is currently an assistant professor of computer science and mathematics at Cadi Ayyad University Marrakech. His main field of research interest is computer security, cryptography, artificial intelligence, intrusion detection and smart cities.

**Mourade Azrour** received his PhD from Faculty of sciences and Technologies, Moulay Ismail University, Errachidia, Morocco. He received his MS in computer and distributed systems from Faculty of Sciences, Ibn Zouhr University, Agadir, Morocco in 2014. Mourade currently works as compter sciences professor at the Department of Computer Science, Faculty of Sciences and Technologies, Moulay Ismail University. His research interests include Authentication protocol, Computer Security, Internet of things, Smart systems. Mourade is member of the member of the scientific committee of numerous international conferences. He is also a reviewer of various scientific journals. Mourade Has edited a scientific book "IoT and Smart Devices for Sustainable Environment" and his is a guest editor in journal "EAI Endorsed Transactions on Internet of Things".

**Said Benkirane** received his PhD from Choaib Dokkali University, El jadida, Morocco in 2013. He is currently a PH professor of computer science and mathematics at Cadi Ayyad University Marrakech. His research interests include computer security, artificial intelligence, smart cities and VANET networks

**Mouaad Mohy-Eddine** received his Master in Computer science and Big Data from Sultan Molay Solaimane University Khouribga, Morocco in 2020. He is currently a PhD student of computer security at Cadi Ayyad University Marrakech. His main field of research interest is machine learning, intrusion detection and IoT security.

**Hanaa Attou** received his engineer diploma in Big Data and Decision making from Mohamed V University, Rabat, Morocco in 2020. She is currently a PhD student of computer security at Cadi Ayyad University Marrakech. His main field of research interest is networking, deep learning and cloud security.

**Maryam Douiba** received his engineer diploma from Hassan I, Settat, Morocco, in 2014. She is currently a PhD student of computer security at Cadi Ayyad University Marrakech. His main field of research interest is machine learning, Blockchain technology and IoT Security.