Hybrid Cyber-Security Model for Attacks Detection Based on Deep and Machine Learning

https://doi.org/10.3991/ijoe.v18i11.33563

Shaymaa Mahmood Naser[1(✉)], Yossra Hussain Ali[1], Dhiya Al-Jumeily OBE[2]
[1] Computer Science Department, University of Technology, Baghdad, Iraq
[2] Faculty of Engineering and Technology, Liverpool John Moores University, UK
cs.19.29@grad.uotechnology.edu.iq

Abstract—Nowadays, numerous attacks can be considered high risks in terms of the security of Wireless Sensor Network (WSN). As a result, different applications are introduced to manage the data and information exchange and related security sides to be saved in the transmission of data. Recently, most of the security attacks are classified as cyber ones. These attacks interest in the system halting and destroying the data rather than stealing the data. In this paper, a cyber-attacks detection system is proposed based on an intelligent hybrid model that uses deep and machine learning technologies. The proposed model improves the cyber-attack detection speed. In addition, a feature reduction model is proposed using machine learning methods (PCA and SVD) to select the most related features to the adopted classes of attacks. This can affect positively the deep-learning model complexity. The obtained results demonstrate the superiority of the proposed hybrid model-based cyber detection system in comparison to the traditional ones in reaching an accuracy of 99.98%, 100%, 100%, 100% for precision, recall, and F1-measure respectively, and reducing the time to 23s for the datasets of Message Queuing Telemetry Transport-Dataset (MQTT-DS) and Wireless Sensor Networks Dataset (WSN-DS).

Keywords—WSN, cyber-attack, deep learning, PCA, SVD, SGD, CNB

1 Introduction

The Wireless Sensor Network (WSN) has been widely used in different applications, which leads to arise the possibility of suffering from a wide range of attacks including cyber ones. Therefore, it is important to propose cyber-security systems that can detect and prevent the attacks that deal with WSN. Cyber-security plays a major role in protecting the framework from threats by adopting artificial intelligence techniques such as Deep-Learning (DL) and Machine-Learning (ML) to build a smart attack detection model in the Internet of Things (IoT) and Wireless Sensor Network (WSN). Artificial Intelligent models require specific cyber-security defense and protection solutions [1][2][3]. The ML and DL methods use numerous algorithms that can be briefly defined. The Principal Component Analysis (PCA) algorithm can be defined as an unsupervised machine learning that is used for dimensionality reduction [4]. The Singular

Value Decomposition (SVD) refers to a matrix factorial that has found several applications for engineering problems [5]. PCA is considered a linear unsupervised feature extraction procedure for minimizing the dimensions of remote sensing data, such as Hyperspectral remote Sensing Images (HSI) [6]. It minimizes a large set of the variables with high correlation to smaller set of uncorrelated variables to be called as principal components. They include the majority of the variations in the original data [7]-[8]. The goal of using SVD is to reduce and convert the high-dimensional space into a low-dimensional domain [9]. The gradient descent method is commonly used in neural network optimization. Gradient descent is a typical method for minimizing an optimization problem's objective function. The number of the required steps that are adopted in detecting a local minimum is defined by the learning rate of a factor. The Stochastic Gradient Descent (SGD) makes frequent updates with a lot of variation, which causes the objective function to change a lot. SGD, on the other hand, is capable of reaching new and potentially superior local minima due to its unpredictability [10]. The Naive Bayes classifier is a simplified version of a Bayesian network, which is a probabilistic model that estimates the probability of categories in each test set using Bayes' theorem. The classifier considers that each feature is conditionally independent of the others, contributing independently but equally to the target's final categorization. Complement Naive Bayes (CNB) Classification Algorithm depend on this concept, it is used to classify find attack or normal flow. Both algorithms are used to show the results of the flow estimation packets [11]. In this paper, an intelligent hybrid model is adopted to design a cyber-attack detection system for different types of them. and utilizing the result with CNB and SGD algorithms gave cyber-attacks detection methods for IoT and WSN. This model use DL and ML methods in three stages feature reduction, feature extraction, and classification. These methods are necessary to detect attacks at an early stage to prevent the loss of available resources. The proposed method is tested over the datasets of MQTT-DS and WSN-DS [12][13]. The obtained results show a promising enhancement in the proposed hybrid-based cyber-security system in comparison with the traditional methods.

2 Related work

Cyber-security is an important field in computing systems that has been investigated by different researchers to produce secured ones that can detect and isolate the cyber-attacks inefficient way, particularly in WSN. In [14], researchers presented the main and sub-categories of ML that were adopted in cyber-security to detect spam, phishing page, malware, DOS attack, and biometric identification. The results proved the claim of the researchers. In [15], many ML techniques were used in cyber-security applications of Vehicular Ad hoc Networks (VANET). This was performed to detect the attack types using the suggested algorithms. These techniques could change the attackers to be ineffective which are simulated using Supervisory Control and Data Acquisition (SCADA) system and snooping discover. While the authors of [16] presented six ML methods, including Nave Bayes, Decision Trees, Random Forests, Neural Network, Gradient Boost, and Multilayer Perception. They were applied to a novel dataset to protect IoT systems. The result proved that the proposed methods were performed efficiently in detecting the attacks. On the other hand, a group of researchers have suggested in [17] deep learning-based correlation prediction models for lightweight subgraph extraction and labeling that used the Weisfeiler-Lehman kernel and Mixed Convolutional Neural Network (WL-DCNN). This was to get better the self-learning ability of topological mining features with a high degree of generality and exceptional performance. The authors of [18] suggested a DL model detect Distributed Denial-of-Service (DDoS) cyber-security attacks on CICIDS2017 datasets with an accuracy of up to 97.16%. In [19], assured and recent research on deep learning already used huge data sets to convolution Neural Networks. This means that today's modern trend is towards intelligence algorithms for deep learning. Finally, in [20], a comparative work and rendering analysis for different machine learning and deep learning methods were proposed for detecting the intruders and saving the WSN-DS dataset. It was shown that deep learning classifiers were the best intrusion detection outcomes than machine learning manners.

3 Proposed system

The proposed cyber-security system is designed based on a hybrid model that adopts deep learning and machine learning mechanisms. Figure 1 shows the structure of the proposed system that includes a multi-stage of workflow that starts from feature reduction and goes through feature extraction till the classification. It is shown from this figure that the input of the proposed system is the adopted datasets of MQTT-dataset (44 feature) [12] and WSN-DS (17 feature) [13]. The preprocessing phase is used for normalizing and reforming the dataset parameters to be ready for input to the feature reduction stage. ML techniques are used in the stage of feature reduction to reduce them from 44 to 15 in the MQTT-DS dataset and from 17 to 15 in WSN-DS. The reduced features are entered into the proposed hybrid model to change the attributes of these features using the power of DL and ML. They are divided into 70% train data and 30% test data to perform the feature extraction model. The last stage uses the ML techniques of SGD and CNB to apply the classifications of the cyber-attacks. Each stage is explained briefly in the next sub-sections.

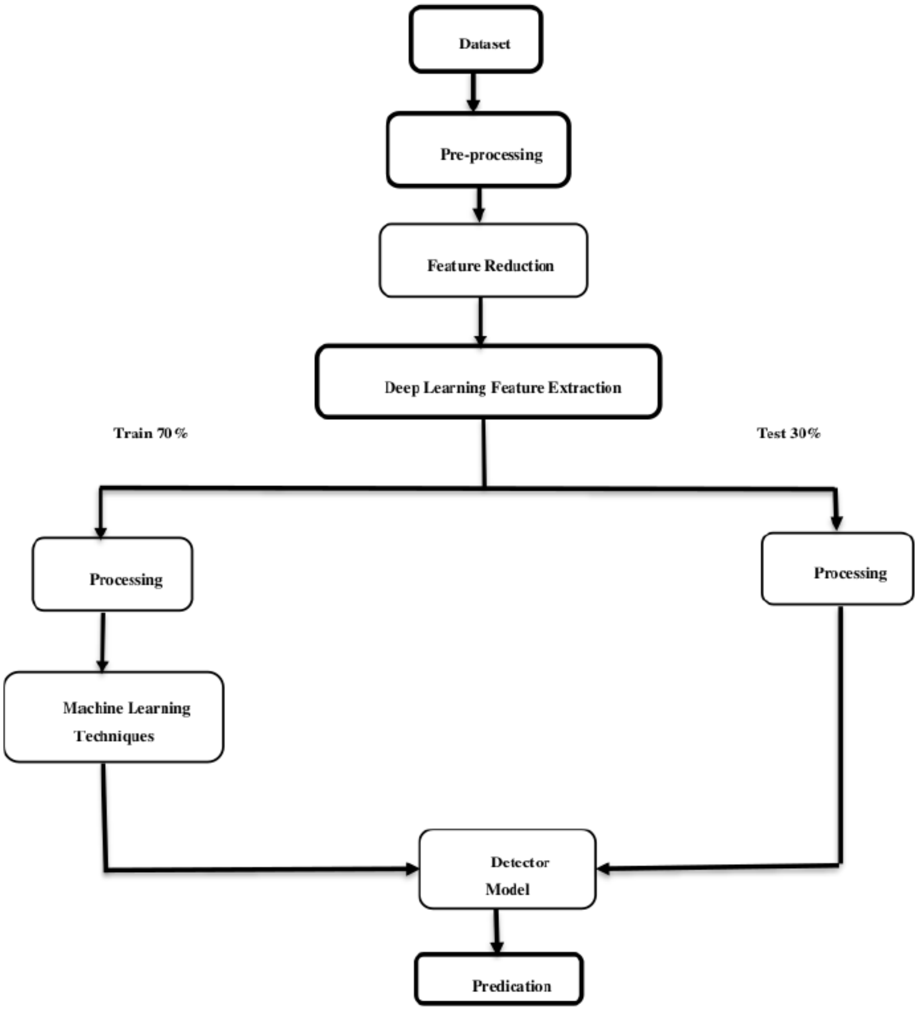

Fig. 1. Proposed system structure

3.1 Pre-processing stage

The adopted datasets are raw data that is gathered from a simulated environment. These datasets need to be handled to ensure that the performance of the proposed system is on a high scale of accuracy. The information of the datasets is disparate types that can be processed by analyzing the data to make misleading results. The main functions in data pre-processing are data cleaning, data integration, data reduction, and data transformation. the data transference consists of various methods including smoothing, attribute construction, aggregation, normalization … etc. Normalization that applies to a dataset in this paper is Min-Max manner, Standard Scaler [21]. The Min-max normalization adopts the linear transformation of the raw data. The $\min_A$ and $\max_A$ refer to the minimum and maximum accounts of features in A. Min-max normalization maps account, vi, of A to v ' i in the range [new $\min_A$, new $\max_A$] by counting [22]:

$$v_i' = \frac{vi - \min_A}{maxA - \min_A}(\text{new } \max_A - \text{new } \min_A) + \text{new } \min_A \qquad (1)$$

Moreover, the StandardScaler is employed in normalizing the accounts by removing the mean and scaling it to unit variance using computing relevant statistics to each attribute. Later on, it saves the mean and standard deviation for them to transform data. The related estimator saves the mean and standard deviation that is computed for the training set. Transforming the unseen test set X is tested in maps that are allocated into the right region of the attribute region [23]:

$$\text{Zscaled} = \frac{(x - \mu)}{\sigma} \qquad (2)$$

where µ = mean, Ơ=standard deviation.

3.2 Feature reduction stage

The reduction of the number of features in a sensing way that ensures keeping the information in the valid formulation is called a feature reduction or dimension reduction. The number of variables is minimized at the time of reducing the number of characteristics. This is for making the computer's tasks easier and faster. Figure 2 shows the block diagram of the adopted feature reduction process. It is clear that this stage uses two ML algorithms of PCA and SVD to minimize the number of features.

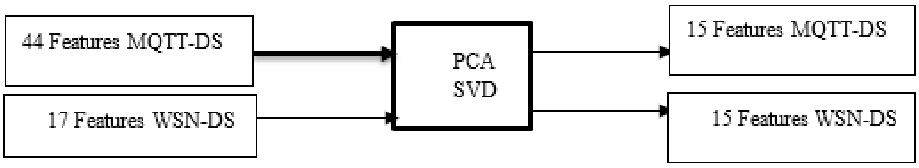

Fig. 2. Feature reduction stage

For more explanation, the produced models of feature reduction using PCA and SVD are briefly illustrated. PCA looks for orthogonal vectors with k dimensions that can be utilized to formulate the data. The original data is thus aimed at a smaller area to reduce the dimensions. PCA combines the attributes in a MQTT-DS and WSN-DS datasets by generating a little size alternate set of variables. The raw data can then be shown to this smaller group. It considers the relationships between attributes that were not suspected in previous states, to be allowed for interpretations that would not normally result. Since PCA works on digital data, data processing has been done before entering this stage [24].

$$C_A = A * \frac{AT}{(n-1)} \qquad (3)$$

SVD is a technique for feature reduction that is formulated as a matrix factorization method. It adopts the Eigen decomposition of a matrix with square dimensions (n x n) to any formula matrix (n x m). SVD method is utilized to decorate the time and to improve the classification performance. The aim of adopting the SVD is to manage the high dimensional space to be a low semantic dimensional space. It also finds the relationship between adopted features [5].

3.3 Proposed deep learning model for feature extraction stage

It should be highlighted that the utmost contribution of this work is to propose a hybrid model that involves the deep-learning structure, which has twenty-seven neural network layers of nine CNN, six MaxPooling, eight LeakyReLU, three Dense, and one Flatten. divided into twelve layers. This hybrid model produces higher predictive performance when compared to the traditional deep learning models alone. Figure 3 shows the deep-learning model, used for feature extraction. Features are selected and then fed to the illustrated architecture. The first convolution with the kernel is built in dimensions size of 3, 16 filters, 1 stride, and (10x1) input shape. The second convolution includes MaxPooling (1 size and 1 stride) and LeakyReLU with alpha 0.3. The third layer contains the convolution with a kernel size of 3, 32 filters, and 1 stride, following MaxPooling (1 size and 1 stride) and LeakyReLU with alpha 0.3. The fourth layer consists of convolution with the kernel size of 3, 64 filters, and 1 stride, also following MaxPooling (1 size and 1 stride) and LeakyReLU with alpha 0.3. The fifth layer has the same formulation as the fourth layer but 128 Dense is added with activation linear. The sixth and seventh layers are similar, while the eighth layer is built with 512 densities for linear activation. Moreover, the ninth and tenth layers contain convolution with 16 filters, kernel size of 3 and padding same, and LeakyReLU alpha of 0.3. Finally, the eleventh layer has a convolution of 35 with the kernel size of 3, stride 1, padding the same, and activation is linear. Then the twelfth layer has a flattened and dense unit of 2 with SoftMax activation function.

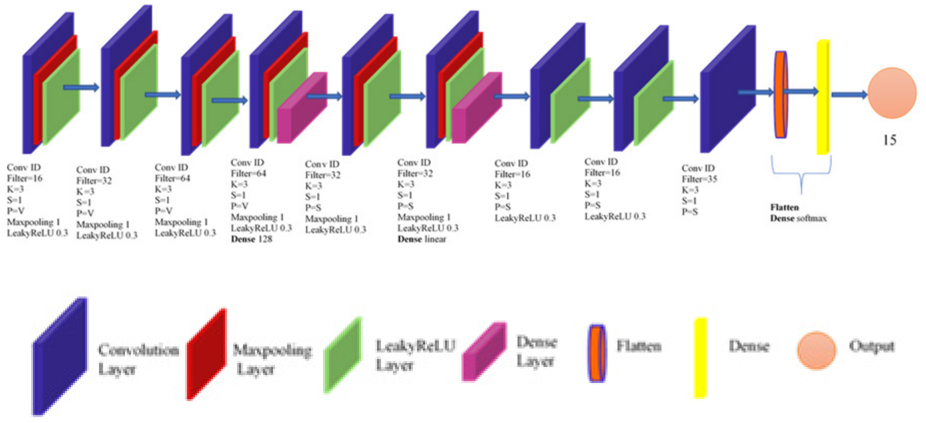

Fig. 3. Deep learning architecture

3.4 Machine learning classification techniques stage

ML classification is considered the most important one in cyber-security systems that generally can model the complex class cases to accept a variety of input predictor data without providing assumptions about the data distribution. In this paper, we use two ML-based classification algorithms Complement Naive Bayes (CNB) and Stochastic Gradient Descent (SGD) to distinguish attacks [25]. CNB adopts the standard Multinomial Naive Bayes algorithm. Multinomial Naive Bayes is not working well in the imbalanced datasets. This type of dataset can cause an overfitting problem in most deep and machine learning models [26]. It is the combination of Bayes' Theorem and the attribute independence assumption; The equation below describes the CNB algorithm.

$$ICNB(d) = argmax_c \left\{ \log p(\theta_c) - \sum_i f_i \log \frac{Nc{\sim}i + ai}{Nc{\sim} + a} \right\} \tag{4}$$

while$N\tilde{c}\,i$ means the number of times attack i occurred in flow in classes other than c, $N\tilde{c}$ means the total number of attack occurrences in classes other than c, while $p(\theta_c)$is the class prior estimate. While SGD is an iterative algorithm that considers a random point as a start level of a function for counting down to achieve the lower pitch of such function. By equaling the slope value to zero, the optimal point can be found by the SGD algorithm. In the case of linear regression, the sum of squared residuals is mapped as the function "y" in terms of the weight vector of "x" [27].

$$\mathrm{w_{t+1}} = \mathrm{w_t} - \gamma_\mathrm{t} \nabla_\mathrm{w} \mathrm{Q}(\mathrm{z_t}, \mathrm{w_t}) \tag{5}$$

where γ is the learning rate sufficiently small, $\{w_t$, t= 1, . . . } depends on the examples randomly picked at each iteration.

4 Performance evaluation criteria

Different criteria are used for measuring the performance of the proposed hybrid algorithms, such as accuracy, precision, recall, and F1-measure. The accuracy that expresses the classification efficiency can be calculated as [28]:

$$\text{Accuracy} = \frac{\text{TN+TP}}{\text{TN+FP+FN+TP}} \tag{6}$$

Where,
TP: Actual class is positive and the algorithm classifies it as positive.
FN: Actual class is positive but the algorithm classifies it as negative.
FP: Actual class is negative but the algorithm classifies it as positive.
TN: Actual class is negative and the algorithms classifies it as negative.

Moreover, the precision that represents the correction ratio of predicting the positive results is evaluated as [28]:

$$\text{Precision} = \frac{\text{TP}}{\text{FP+TP}} \tag{7}$$

On the other hand, the recall, which is the correct prediction of the actual positive results, can be computed as [28]:

$$\text{Recall} = \frac{\text{TP}}{\text{FN+TP}} \tag{8}$$

While the F1-measure is the harmonic mean of recall and precision is calculated as [28] [29]:

$$\text{F1} - \text{measure} = \frac{2 \text{ x Precision x Recall}}{\text{Precision+Recall}} \tag{9}$$

5 Results

It is well known that the proposed systems should be tested over different circumstances to ensure efficiency and prove the claim of the authors. The proposed cyber-security system is tested using the mentioned datasets of MQTT-DS and WSN-DS [12] and [13]. The testing of the efficiency is based on the highlighted parameters of accuracy, precision, recall, F1-measure, and complexity in terms of execution time. The following system is considered for obtaining the performance parameters;

- ML for the system of [12].
- DL for the general proposed deep-learning model,
- PCA15-DL for feature reduction to 15 based on PCA in combination with deep-learning classification.
- SVD15-DL for feature reduction to 15 based on SVD in combination with deep-learning classification.
- HYP-PCA15-DL-CNB for the proposed hybrid system with PCA feature reduction to 15 in combination with the deep-learning model for feature extraction and machine-learning of CNB for attack classification.
- HYP-PCA15-DL-SGD for the proposed hybrid system with PCA feature reduction to 15 in combination with the deep-learning model for feature extraction and machine-learning of SGD for attack classification.
- HYP-SVD15-DL-CNB for the proposed hybrid system with SVD feature reduction to 15 in combination with the deep-learning model for feature extraction and machine-learning of CNB for attack classification.
- HYP-SVD15-DL-SGD for the proposed hybrid system with SVD feature reduction to 15 in combination with the deep-learning model for feature extraction and machine-learning of SGD for attack classification. Table 1 explains the first indicator of performance efficiency of the proposed system, which is accuracy. It is shown that the accuracy of the proposed hybrid cyber-security systems keeps the close range to the ML [12] and other compared systems. The proposed system gains more accuracy for WSN-DS [13] dataset. The profit of this system is clearly shown in the next tables of results.

Table 2 illustrates the recall ratio for the compared eight systems. The proposed hybrid system records a high recall ratio of 100% to gain a significant improvement in comparison with other compared systems for both datasets. This refers to the addressing of the right prediction of the attacks and the features that are extracted from deep learning are in place. The ML classification of the hybrid systems also supports the reduction of overfitting and redundant feature occurring. Table 3 shows the enhancement of the precision parameters for most of the proposed systems, particularly the hybrid ones that record 100%. This indicator explains the benefit of using the deep-learning models as well as the combination of ML and DL in both feature extraction and classification. Table 4 records another superior performance of the proposed hybrid systems in comparison with the traditional ones. A significant improvement is measured for the hybrid system to reach 100% in the F1 measure due to the use of the magic combination of DL and ML as well as the feature reduction. The features are reduced in the first stage and the overfitting is almost disappeared, while the use of DL feature extraction considers the similarity in the distinct input features, in which the overfitting is disappeared completely. On the other hand, Table 5 expresses the complexity of the compared eight systems in terms of the required time for execution. It is important to note that the execution time of the ML [12] is equal to 4100 seconds and the feature reduction with the deep-learning (PCA15-DL) takes 3800 seconds for 100 epochs, while the proposed hybrid systems consume just 23 seconds to get the promising results explained in Tables 1, 2, 3, and 4. This notified reduction in the time is returned to the need of just one epoch to obtain the required results with precision, recall, F1-measure of 100%, and accuracy up to 99.9%. This is a significant improvement recorded to the proposed hybrid cyber-security that can enhance all performance parameters with a very low consumed time, which reflects positively on the activation of different applications.

Table 1. Performance accuracy of the eight compared systems

Dataset	**ML**	**DL**	**PCA15-DL**	**SVD15-DL**	**HYP-PCA15-DL-CNB**	**HYP-PCA15-DL-SGD**	**HYP-SVD15-DL-CNB**	**HYP-SVD15-DL-SGD**
MQTT-DS-Biflow (bidirectional)	96.61% - 99.97%	99.85%	99.91%	99.78%	99.74%	99.74%	99.73%	99.74%
MQTT-DS-Un flow(unidirectional)	78%-99.98%	99.79%	99.8%	99.87%	99.78%	99.78%	99.76%	99.79%
MQTT-DS-Packet	65.39%-88.55%	99.99%	99.77%	99.86%	99.98%	99.98%	99.98%	99.98%
WSN-DS	75.6%-99.4%	97.9%	98.04%	98.49%	99.98%	99.98%	99.98%	99.98%

Table 2. Performance Recall of the eight compared systems

Dataset	**ML**	**DL**	**PCA15-DL**	**SVD15-DL**	**HYP-PCA15-DL-CNB**	**HYP-PCA15-DL-SGD**	**HYP-SVD15-DL-CNB**	**HYP-SVD15-DL-SGD**
MQTT-DS-Biflow (bidirectional)	96.61% - 99.9%	27.13%	27.21%	27.41%	100%	100%	100%	100%
MQTT-DS-Uniflow(unidirectional)	78%-99.98%	24%	24%	24%	100%	100%	100%	100%
MQTT-DS-Packet	65.39%-88.55%	99.98%	24%	24%	100%	100%	100%	100%
WSN-DS	97.2%-98.5%	97.89%	98%	98.49%	100%	100%	100%	100%

Table 3. Performance Precision of the eight compared systems

Dataset	**ML**	**DL**	**PCA15-DL**	**SVD15-DL**	**HYP-PCA15-DL-CNB**	**HYP-PCA15-DL-SGD**	**HYP-SVD15-DL-CNB**	**HYP-SVD15-DL-SGD**
MQTT-DS-Biflow (bidirectional)	97.02% - 99.9%	100%	100%	100%	100%	100%	100%	100%
MQTT-DS-Uniflow(unidirectional)	88.91% - 99.98%	100%	100%	100%	100%	100%	100%	100%
MQTT-DS-Packet	65.42%-88.55%	100%	100%	100%	100%	100%	100%	100%
WSN-DS	97.4% - 98.7%	97.86%	98.07%	98.5%	100%	100%	100%	100%

Table 4. Performance F1-measure of the eight compared systems

Dataset	**ML**	**DL**	**PCA15-DL**	**SVD15-DL**	**HYP-PCA15-DL-CNB**	**HYP-PCA15-DL-SGD**	**HYP-SVD15-DL-CNB**	**HYP-SVD15-DL-SGD**
MQTT-DS-Biflow (bidirectional)	96.15% - 99.9%	42.68%	42.78%	43.03%	100%	100%	100%	100%
MQTT-DS-Uniflow(unidirectional)	75.26%-99.98%	38.74%	38.72%	38.83%	100%	100%	100%	100%
MQTT-DS-Packet	60.4% - 88.54%	99.99%	38.74%	38.81%	100%	100%	100%	100%
WSN-DS	-	97.85%	98.05%	98.49%	100%	100%	100%	100%

Table 5. Performance complexity in terms of execution time

	DL For 100 epochs (sec)	**PCA15-DL For 100 epochs (sec)**	**HYP-PCA15-DL-CNB** **HYP-PCA15-DL-SGD** **HYP-SVD15-DL-CNB** **HYP-SVD15-DL-SGD**
Time	4100	3800	23 second

6 Conclusion

An Intelligent hybrid cyber-security system was proposed for detecting and preventing the related attacks under the WSN environments. This system adopted the ML SVD and PCA for feature reduction and efficiently reducing the dataset features. It also adopted the DL model for further feature extraction and ML SGD and CNB for attack classification. The combination of the ML and DL provided the system with early detection and learning systems with high-performance parameters that indicated efficiency. Five performance parameters were adopted: accuracy, precision, recall, F1-measure, and execution time as complexity. The achieved results showed the superior performance of the proposed hybrid system with different DL and ML techniques in terms of the five performance parameters. The accuracy was enhanced in WSN-DS and kept at the same level for the MQTT-DS. While the significant enhancements in the precision, recall, and F1-measure to reach 100%. The execution time is reduced significantly due to the need for just one epoch to reach the required performance in terms of training and detection of cyber-attacks.

7 Acknowledgment

This work was supported by the University of Technology.

8 References

[1] Kahina CHELLI. (2015). Security Issues in Wireless Sensor Networks: Attacks and Countermeasures. Proceedings of the World Congress on Engineering. Volume 1.

[2] Daojing He, Sammy Chan, and Mohsen Guizani. (2017). Cyber Security Analysis Protection of Wireless Sensor Networks for Smart Grid Monitoring. IEEE Wireless Communications. Volume 24. Issue 6. pp 98 – 103. https://doi.org/10.1109/MWC.2017.1600283WC

[3] Jian-hua LI. (2018). Cyber security meets artificial intelligence: a survey. Frontiers of Information Technology & Electronic Engineering. volume 19. pp1462–1474. https://doi.org/10.1631/FITEE.1800573

[4] N. Alseelawi, and H. T. Hazim. (2022). A Novel Method of Multimodal Medical Image Fusion Based on Hybrid Approach of NSCT and DTCWT. International Journal of Online Biomedical Engineering. volume 18. issue 3. https://doi.org/10.3991/ijoe.v18i03.28011

[5] Prashant Johri, Jitendra Kumar Verma, Sudip Paul. eBook, Springer, Nature Singapore : Applications of Machine Learning", Algorithms for Intelligent Systems. (2020). https://doi.org/10.1007/978-981-15-3357-0

[6] Palash Uddin, Al Mamun, Ali Hossain. (2020). PCA-based Feature Reduction for Hyperspectral Remote Sensing Image Classification. IETE Technical Review. Volume 38. Issue 4. https://doi.org/10.1080/02564602.2020.1740615

[7] Priyankaa ,Dharmender Kumar. (2020). Feature Extraction and Selection of kidney Ultra sound Images Using GLCM and PCA. International Conference on Computational Intelligence and Data Science. Procedia Computer Science. https://doi.org/10.1016/j.procs.2020.03.382

[8] Farzana Anowar, Samira Sadaoui, Bassant Selim. (2021). Conceptual and empirical comparison of dimensionality reduction algorithms (PCA, KPCA, LDA, MDS, SVD, LLE, ISOMAP, LE, ICA, t-SNE). Computer Science Review. Volume 40. https://doi.org/10.1016/j.cosrev.2021.100378

[9] Razieh Rastgoo, Kourosh Kiani, Sergio Escalera. (2021). Real-time isolated hand signlan gauge recognition using deep networks and SVD. Journal of Ambient Intelligence and Humanized Computing. volume 13, p.p591–611. https://doi.org/10.1007/s12652-021-02920-8

[10] Mert Alagözlü. (2020). Gradient Descent and Stochastic Gradient Descent.

[11] Anthony Kelly, Marc Anthony Johnson. (2021). Investigating the Statistical Assumptions of Na¨ıve Bayes Classifiers. 2021 55th Annual Conference on Information Sciences and Systems (CISS). IEEE. https://doi.org/10.1109/CISS50987.2021.9400215

[12] Hanan Hindy, Miroslav Bures, Christos Tachtatzis, Xavier Bellekens. (2020). Machine Learning Based IoT Intrusion Detection System: An MQTT Case Study. ArXiv. Springer.

[13] Iman Almomani, Bassam Al-Kasasbeh, Mousa AL-Akhras. (2016). WSN-DS: A Dataset for Intrusion Detection Systems in Wireless Sensor Networks. Journal of Sensors. Hindawi Publishing Corporation. Volume 2016. https://doi.org/10.1155/2016/4731953

[14] Tony Thomas Athira, P. Vijayaraghavan, Sabu Emmanuel. (2020). Machine Learning Approaches in Cyber Security Analytics. Springer Nature Singapore. https://doi.org/10.1007/978-981-15-1706-8

[15] Anand Handa, Ashu Sharma, Sandeep K. Shukla. (2019). Machine learning in cybersecurity: A review. Wiley. https://doi.org/10.1002/widm.1306

[16] I. Vaccari, G. Chiola, M. Aiello, M. Mongelli, and E. Cambiaso. (2020). MQTTset, a New Dataset for Machine Learning Techniques on MQTT. Sensors. MDPI. https://doi.org/10.3390/s20226578

[17] Ru Huang, Lei Ma, Guangtao Zhai, Jianhua He, Xiaoli Chu, Huaicheng Yan. (2020). Resilient Routing Mechanism for Wireless Sensor Networks with Deep Learning Link Reliability Prediction. IEEE Access. Volume 8. pp 64857 – 64872. https://doi.org/10.1109/ACCESS.2020.2984593

[18] Monika Roopak, Gui Yun Tian, Jonathon Chambers. (2019). Deep Learning Models for Cyber Security in IoT Networks. 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC). IEEE. https://doi.org/10.1109/CCWC.2019.8666588

[19] H. Salim, S. H. Abbood, M. S. d Rahim, and A. M.Alaidi. (2022). DR-LL Gan: Diabetic Retinopathy lesions synthesis using Generative Adversarial Network. International journal of online and biomedical engineering, volume 18. issue 3. https://doi.org/10.3991/ijoe.v18i03.28005

[20] Pankaj R. Chandre, Parikshit N. Mahalle and Gitanjali R. Shinde. (2020). Deep Learning and Machine Learning Techniques for Intrusion Detection and Prevention in Wireless Sensor Networks: Comparative Study and Performance Analysis. Design Frameworks for Wireless Networks. Lecture Notes in Networks and Systems 82. https://doi.org/10.1007/978-981-13-9574-1_5

[21] Lars Buitinck, Gilles Louppe, Mathieu Blondel , Fabian Pedregosa, Andreas C. M¨uller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake Vanderplas, Arnaud Joly, Brian Holt, Ga¨el Varoquaux. (2013). API design for machine learning software: experiences from the scikit-learn project. arXiv. European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases.

[22] Jiawei Han,Micheline Kamber, Jian Pei. Data Mining Concepts and Techniques. Third Edition. 2012. Elsevier.

[23] V N Ganapathi Raju; K Prasanna Lakshmi; Vinod Mahesh Jain; Archana Kalidindi; V Padma. (2020). Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification. 2020 Third International Conference on Smart System and Inventive Technology (ICSSIT). IEEE. https://doi.org/10.1109/ICSSIT48917.2020.9214160

[24] Karunakar Pothuganti. (2020). Overview on Principal Component Analysis Algorithm in Machine Learning. International Research Journal of Modernization in Engineering Technology and Science. Volume 02. Issue10.

[25] Aaron E. Maxwell, Timothy A. Warner, Fang Fang. (2017). Implementation of machine-learning classification in remote sensing: an applied review. International Journal of Remote Sensing. Volume 39. issue 9 pp 2784-2817. https://doi.org/10.1080/01431161.2018.1433343

[26] Berna Seref, Erkan Bostanci. (2019). Performance of Naïve and Complement Naïve Bayes Algorithms Based on Accuracy, Precision and Recall Performance Evaluation Criterions. International Journal of Computing Academic Research (IJCAR). Volume 8. issue 5. pp 75-92.

[27] L´eon Bottou. (2012). Stochastic Gradient Descent Tricks. Part of Lecture Notes in Computer Science book serires. Volume 7700. Neural Networks: Tricks of Trade. Pp 421-436. https://doi.org/10.1007/978-3-642-35289-8_25

[28] H. T. ALRikabi and H. T. Hazim. (2021). Enhanced Data Security of Communication System Using Combined Encryption and Steganography. International Journal of Interactive Mobile Technologies, volume 15. issue 16. https://doi.org/10.3991/ijim.v15i16.24557

[29] Berna Seref, Erkan Bostanci. (2018). Sentiment Analysis using Naive Bayes and Complement Naive Bayes Classifier Algorithms on Hadoop Framework. 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT). IEEE. https://doi.org/10.1109/ISMSIT.2018.8567243

9 Authors

Shaymaa Mahmood Naser Programmer is presently one employee of Federal Board of Supreme Audit Editorial in Al baghdad, Iraq. He received his B.Sc. degree in computer science in 2002 from the Al-Mustansiriya University in Baghdad, Iraq. His M.Sc. degree in computer science focusing on security of wireless sensor networks from Information Institute for Postgraduate Studies in Baghdad/Iraq in 2017. His current research interests include intelligent techniques in cyber-security system for wireless sensor networks and the Internet of Things (IoT). Baghdad City-Al Otayfia, Iraq (e-mail: cs.19.29@grad.uotechnology.edu.iq). The number of articles in security of wireless sensor networks – 3, The number of articles in cyber security – 3.

Yossra Hussain Ali Assistant Professor. She received her B.Sc., M.Sc., and Ph.D. degrees in 1996, 2002, and 2006 respectively in Computer Sciences from Iraq, University of technology, department of Computer Sciences. She joined the University of Technology, Iraq in 1997. During her postgraduate studies, she worked on Computer networks, Information systems, Agent Programming, and Image Processing, she has some experience in Artificial Intelligent and Computer Data Security, She Reviewer at many conferences and journals, and she supervised undergraduate and postgraduate (Ph.D. and M.Sc.) dissertations for many students in Computer sciences, she has several professional certificates, Yossra has published in well-regarded journals where she published 65 papers in different journals and conferences (email: yossra.h.ali@uotechnology.edu.iq).

Dhiya Al-Jumeily OBE is a professor of Artificial Intelligence and the president of eSystems Engineering Society. Dhiya is the founder and general series chair of the IEEE International Conference on Developments in eSystems Engineering DeSE (dese.org.uk) since 2007. He is a Senior Member of the IEEE and a Chartered IT Professional. He is also a fellow of the UK Higher Education Academy. 2019, Kazan Federal University, Russia, Visiting Professor, in 2000, Liverpool John Moores University, England, PhD, Intelligent Software Development, in 1995, Liverpool John Moores University, United Kingdom, MPhil in Computer Science. Where in 1992, University of Liverpool, United Kingdom, MSc, Applied Mathematics.1988, University of Baghdad, Iraq, BSc: Mathematics.Technical University of Baghad, Iraq, Lead University Research Advisor. He has extensive research interests covering a wide variety of interdisciplinary perspectives concerning the theory and practice of Applied Artificial Intelligence in medicine, human biology, environment, intelligent community and healthcare. He has published well over 300 peer reviewed scientific international publications, 17 books and 17 book chapters, in multidisciplinary research areas including: Machine Learning, Neural Networks, Signal Prediction, Telecommunication Fraud Detection, AI-based clinical decision-making, medical knowledge engineering, Human-Machine Interaction, intelligent medical information systems, sensors and robotics, wearable and intelligent devices and instruments (email: D.Aljumeily@ljmu.ac.uk).

Article submitted 2022-06-23. Resubmitted 2022-07-27. Final acceptance 2022-07-29. Final version published as submitted by the authors.