AN ENHANCED FEATURE SELECTION AND CLASSIFICATION MODEL FOR NETWORK INTRUSION DETECTION SYSTEM USING DATA MINING TECHNIQUES

Olaiya Folorunsho

Postdoctoral Research Fellow, Unit for Data Science and Computing
North-West University, Potchefstroom Campus, South Africa
&
Lecturer, Department of Computer Science, Federal University Oye Ekiti
Oye Ekiti, Ekiti State, Nigeria
olaiya.folorunsho@fuoye.edu.ng

Ismail Adelabu Adegbola

Head of Department, Department of Computer Science, Oyo State College of Education
Lanlate, Oyo State, Nigeria
ismailadegbolaa@gmail.com

Rasheed Gbenga Jimoh

Professor, Department of Computer Science, University of Ilorin
Ilorin, Kwara State, Nigeria
Jimohrasheed@yahoo.com

Abstract

Security of information in this Information Technology (IT) era has been one of the challenges facing individuals and organisations. Among the measures developed by security experts to counter security threats is the Intrusion Detection System (IDS). Despite earlier research efforts to develop formidable IDSs, the existing systems still suffer from a high false alarm and inability to detect new (novel) attacks because of the high volume of features in network traffic. Therefore, this study aimed at developing IDS with an enhanced feature selection and classification method using two stages of attack identification. The feature selection phase employed Particle Swarm Optimization (PSO) to optimally select relevant features from Principal Component Analysis (PCA)'s projected principal space. The reduced dataset was passed into the misuse detector using C4.5 to classify network traffic into normal and attack. The "assumed" normal traffic further passed to the anomaly detector, the second-level classifier using Support Vector Machine (SVM) for detecting new attacks that the misuse detector has not previously detected. The proposed model was demonstrated on the KDD Cup'99 and NSL-KDD intrusion datasets, with the system achieving a false alarm rate of 0.53% and detection rate of 99.43% for NSL KDD dataset. The results show that enhancing the feature selection phase and classification method reduces the false alarm and improves the system's ability to detect zero-day attacks.

Keywords: **Network traffic; Feature selection; Classification; Data Mining.**

1. Introduction

Over the recent decades, both internet usage and data shared have continued to increase geometrically. This development has conversely revealed numerous security threat issues to information passed across this medium. Intruders are becoming highly sophisticated in their deals to breaking the protective mechanism put in place to exploit the network issue [Di Mauro *et al*.,(2021); Folorunsho & Jimoh,(2017); Ogonji *et al*.,(2020)]. The concept of intrusion emanated through a paper delivered by [Anderson, (1980)], which stated that the users' logs and records of the computers enable the computer to be monitored so easily. It aimed to have strong protections for the information accessed and guided against authorised users and misuse of privilege exclusively reserved for legitimate users. So, this leads to the beginning of the development of host-based IDS, and different researchers have engaged themselves in the designs that lead to improving systems using various methods.

The IDS has been of great importance in the Information Security (IS) domain because of its prompt detection of attacks within a corporate network. Generally, IDS raises the alarm by notifying network administrators whenever security violations are detected. Such infringement could be an aftereffect of break-in endeavour by authorised external users attempting to gain access to the system illegally, thereby compromising it [Eskandari *et al*.,(2020); Kumari & Sharma,(2018)]. Unlike some other security measures, Active IDS will stop the attack and take various steps to tackle it and add to its knowledge base the malicious activities for future reoccurrences. So, it does not give room for an attacker to deny the attack or remove evidence of attack [Folorunsho & Jimoh,(2017)].

Two main design approaches are used in building IDS: misuse and anomaly methods. Misuse detection is developed using the profile of well-defined attacks and to search for the incidence of attacks from the in-coming traffics are compared against the knowledge base of the attacks' signature, and any match indicates attack [Agrawal & Tapaswi,(2019)]. On the other hand, anomaly detection is built using the expected usage pattern of the legitimate activities and the pattern deviating from the normal is considered as an attack [Fernandes et al.,(2019); Moustafa *et al*.,(2017)].

Misuse detection can effectively detect known attacks with a high degree of accuracy and relatively low false alarm, making it have higher accuracy when compared with the anomaly approach over the same attacks. The major shortcoming of misuse detection is the inability to detect unknown (novel) attacks, and also getting enough data on known attacks and currently updating them with the new attacks with the assistance of experts [Zarpelao *et al*.,(2017); Liu & Lang, (2019); Khraisat *et al*.,(2020)]. Moreover, the anomaly approach can accurately detect a zero-day (novel) attack with a high false alarm [Manzoor *et al*.,(2017)]. The hybrid IDS was proposed to perfect the shortcomings of both the misuse and anomaly detections and keep their advantages [Pradhan *et al*.,(2020); Bovenzi *et al*.,(2020); Mohammadi & Amiri,(2019)]. Along these lines, the system benefits from a high detection rate on known attacks in the misuse approach and detecting the unknown (novel) attacks in the anomaly approach [Di Mauro *et al*.,(2021); Moustafa *et al*.,(2017); Manzoor *et al*.,(2017)].

Researches in the domain of IDS have adopted several techniques for its development so that effort could be made to minimise the level of intrusion in the system. Such techniques include honey pot, multiclass, statistical method, multi-agent, mobile agent, artificial intelligence and data mining. The IDS has been automated through the development processes via neural network [Chandak *et al*.,(2019); Ramakrishnan & Devaraju,(2017)], fuzzy inference system [Hider *et al*.,(2017); Stehlik *et al*.,(2017)], Evolutionary Computation [Xue *et al*.,(2018); Gu *et al*.,(2019)], Support Vector Machine [Krishnaveni *et al*.,(2020); Ingre *et al*.,2017)], Decision Tree [Rathore *et al*.,(2018); Mirsky *et al*.,(2018)], and so on.

Regardless of the intrinsic capability of the hybrid intrusion detection system, three imperative issues hinder its overall performance. First, the performance of the system depends on labeled and time-to-time training set with various types of attacks [Maglaras & Jiang,(2014)]. Secondly, the volume of data collected and stored in today's database surpasses the capacity to be directly analysed without using automated analysis techniques [Juumi *et al*.,(2017)]. Lastly, the productive and powerful combination of various detection technologies possesses daunting hindrances when building practical and operational hybrid intrusion detection systems. These constraints have prompted an expanding interest in research in information security in developing IDSs in recent years using data mining techniques to face the difficulties associated with misuse detection techniques. Data Mining is the process that automatically searches through a large dataset to generate and extract implicit and potential patterns using combined methods from statistics, machine learning, and so on [Al-hamami & Alawneh,(2012)]. Conversely, these have brought about variously supervised and unsupervised machine learning techniques for IDS designs.

2. Literature Review

[Sheen and Rajesh,(2008)] presented IDS using C4.5 decision tree algorithm. Filter approach was used in the selection stage and the C4.5 for the classification. The FS comprised three algorithms: ReliefF, IG and chi-square. The performance of the three FS algorithms was evaluated against 20 subset features. The results of the filter approaches showed that the performance of chi-square and IG out-performed that of ReliefF, and the accuracy rate of the ReliefF, IG and chi-square were 95.64%, 95.85% and 95.85, respectively. [Ektefa et al., (2010)] presented IDS using data mining techniques to classify network traffic into normal and attack. In their work, two different classifiers were utilised; C4.5 and SVM. The experiments were carried out on KDD Cup'99 to predict the dataset into the attack and normal. The two classifiers were compared against KDD Cup'99 to determine a better algorithm. The DR for C4.5 and SVM were 84.15% and 83.77%, respectively, while the FAR of C4.5 and SVM were 0.81% and 1.65%, respectively. The results of the comparative analysis showed that C4.5 out-performed SVM.

[Mohammed and Awadekarimi,(2011)] proposed hybrid IDS using a data mining technique. The misuse detector was modeled using ID3 and the distance-based clustering for the anomaly detector. Experiments were performed on the DARPA dataset and the real-life data. In their work, the anomaly detector utilised cannot be evaluated due to the low performance of the classifier on large datasets. The experimental results obtained showed that the data mining-based IDS has an acceptable level of false alarms. [Ahmad *et al*., (2011)] proposed intrusion detection using MLP algorithm. They proposed a combination of feature reduction using PCA and

optimal feature selection using GA. The PCA was used to project component features into the principal scape based on the values of eigenvalues. The GA was used to optimally select relevant features which have high sensitivity for the MLP classifier. The series of experiments on KDD Cup'99 showed that when the PCA_GA feature selection technique was adopted, the original dataset reduced from 41 features to 12 features. As a result of the reduction in the size of the features, the overheads incurred in the computation were reduced and the accuracy rate of the system was improved.

[Chung and Wahid,(2012)] presented IDS using a hybridised Simplified Swarm Optimization (SSO) and weight local search (WLS). The Intelligent Dynamic Swarm (IDS) and rough set (RS) were used for the feature selection stage. Weight local Search was incorporated in the SSO to improve the performance of the classifier. A series of experiments were carried out on KDD Cup'99 to determine the comparative performance of the hybrid system against other classifiers. The accuracy rate of SSO-WLS, SSO, PSO, NSW, and SVM were 93%, 89.6%, 92.1% and 86.8%, respectively. Based on the performance of the system, it showed that hybridised SSO-WLS is one of the competitive classifiers that can be used in IDS development. [Nadiammai and Hemalatha, (2013)] presented data mining-based techniques for hybrid IDS using the combination of snort based statistical and semi-supervised classification algorithm was used in training 2500 data instances. The experimental results were conducted using KDD Cup'99 dataset. The proposed hybrid IDS detect 149 attacks from 180 attacks. The experimental results showed that the hybrid IDS had FAR and an accuracy rate of 2.53% and 98.8%.

[Elngar *et al*.,(2013)] presented a real-time anomaly IDS. PSO and Information Entropy Minimization (IEM) were utilised in the feature selection stage. Hidden Naïve Bayes (HNB) was used as the classifier. The performance of the classifier was evaluated on the intrusion dataset NSL-KDD. The FS method adopted reduced the 14 features in the data set to 13 features. The results of FS were compared with that of PCA and gain ratio. The experimental results showed that the IDS had an accuracy rate of 98.2%. [Subba *et al*.,(2016)] presented an intrusion detection system that was Neural Network-based. PCA was utilised for dimensionality reduction for the feature selection stage, while ANN was used to classify the traffic. Their model utilised both feed-forward and the back-propagation algorithms and some other optimisation techniques to reduce the overall computational overhead and maintain a high-performance level. The experimental results were evaluated using NSL-KDD dataset. The results showed that the ANN-based IDS model had an accuracy and detection rate of 98.86% and 95.77%, respectively.

[Tesfashun and Bhaskari,(2017)] proposed a hybrid IDS that combines the misuse detector to detect the known attacks with the anomaly detector to detect the zero-day attacks. The misuse detector is modeled using random forests classifier while the anomaly detector was modeled using SVM. Their system was simulated using WEKA and Matlab 2013 and evaluated using the NSL-KDD data set. The performance evaluation of their IDS showed that their system had a false positive rate and the attack detection rate of 6.42% and 92.13% respectively. [Shenfield et al.,(2018)]proposed an intelligent intrusion detection system using artificial neural networks. The misuse detector is modeled using the random forests algorithm while the anomaly detector uses multi-layer perceptron. The IDS was implemented to predict network traffic into normal and attack and implemented using MATLAB (2016b). The evaluation carried out on their developed system showed an accuracy of 98%.

[Soheily-Khah et al.,(2018)] presented a hybrid IDS using data mining. They combined both supervised and unsupervised learning using k-means and random forest algorithms. A series of experiments were carried out on the ISCX dataset to evaluate the performance of their system. The system's AC rate, DR rate, and FAR were compared with SVM, Naïve Bayes, 1-NN, Decision Tree Neural Network, and Random Forest algorithms. The results of their developed IDS using kM-RF showed that the DR and FAR was 99.19%, 9.15% and 0.12% respectively. [Salo *et al*.,(2019)] presented a hybrid dimensionality reduction using the combination of information gain (IG) and PCA. The comparative analysis of their model was carried out on three different intrusion datasets; NSL-KDD, ISCX 2012, and Kyoto 2006 using four classifiers; SVM, instance-based K-nearest neighbour (IBK), MLP, and Ensemble. The experimental results showed that Ensemble classifier significantly out-performed the other three classifiers on ISCX 2012 dataset with an accuracy rate of 99.01 and FAR of 1.00%.

[Moustafa *et al*.,(2018)] presented IDS using correlation coefficient for feature selection and Ensemble classifier comparison decision tree (DT), Naïve Bayes and ANN for classification for detecting Denial of Service (DoS) attacks. The experimental results on UNSW-NB 15 dataset had an accuracy of 98.54% for detecting DoS using ensemble classifier. [Nimbalkar and Kshirsagar,(2021)]. presented feature selection for IDS in the Internet-of-Things. Information Gain (IG) and Gain Ratio (GR) were used for the feature selection for detecting DoS and DDoS attacks. The system developed obtained feature subsets using union and insertion operations by ranking top 50% GR and IG features. The experiment was performed on the IoT-BoT and LDD Cup'99 dataset using the reduced datasets of 16 and 19 respectively. The comparative analysis of their system showed that the accuracy rate of Bot-IoT dataset out performed that of KDD Cup'99 dataset.

The literature review observed that NSL-KDD, UNSW-NB 15 and KDD Cupp'99 were the majorly used intrusion dataset. The findings from the literature showed that although PCA was used mostly by the researchers for feature selection, but the IDSs still have a relatively high false alarm and the low detection of novel attacks.

This paper proposes an optimised feature selection and the classification method for the detection of network intrusion attacks with low false alarm and detection accuracy using two-level classification.

3. Method

This section proposes the development of an IDS using soft computing. It starts by describing the system architecture of the proposed model, the algorithms and the various stages involved are presented. The architecture of the proposed system was strictly followed.

3.1. System Architecture

The block diagram of the IDS is presented in Fig.1 which corresponds to the various stages of the system development: data preprocessing module which involves the data cleaning and normalization, feature reduction and selection module which involves the optimization of Principal Component Analysis using Particle Swarm Optimization, classification module which involves misuse detector and anomaly detector.

3.2 Description of Data Sets

This work utilized NSL-KDD and KDD Cup'99 datasets from the University of California Irvine (UCI) machine learning repository. NSL-KDD is an improved version of KDD Cup'99 collected from UCI. It comprises of selected records of the complete KDD dataset and have some advantages over the KDD dataset; it does not include redundant records in the train set, the number of selected records from each difficulty level group is inversely proportional to the percentage of records in the original KDD data set, the number of records in the train and test sets is reasonable which makes it affordable to run the experiments on the complete set without the need to randomly select a small portion. This data set is meant to solve the problems associated with the KDD Cup'99 data set. Each instance is a series of TCP packets recorded within a well-defined time. It is made up of 42 nominal and continuous features including one label representing the class. The NSL-KDD dataset of twenty-five thousand, one hundred and ninety-two (25192) instances were selected at random from its dataset. In the KDD Cup'99 dataset, twenty-five thousand and eight (25008) instances were randomly selected.

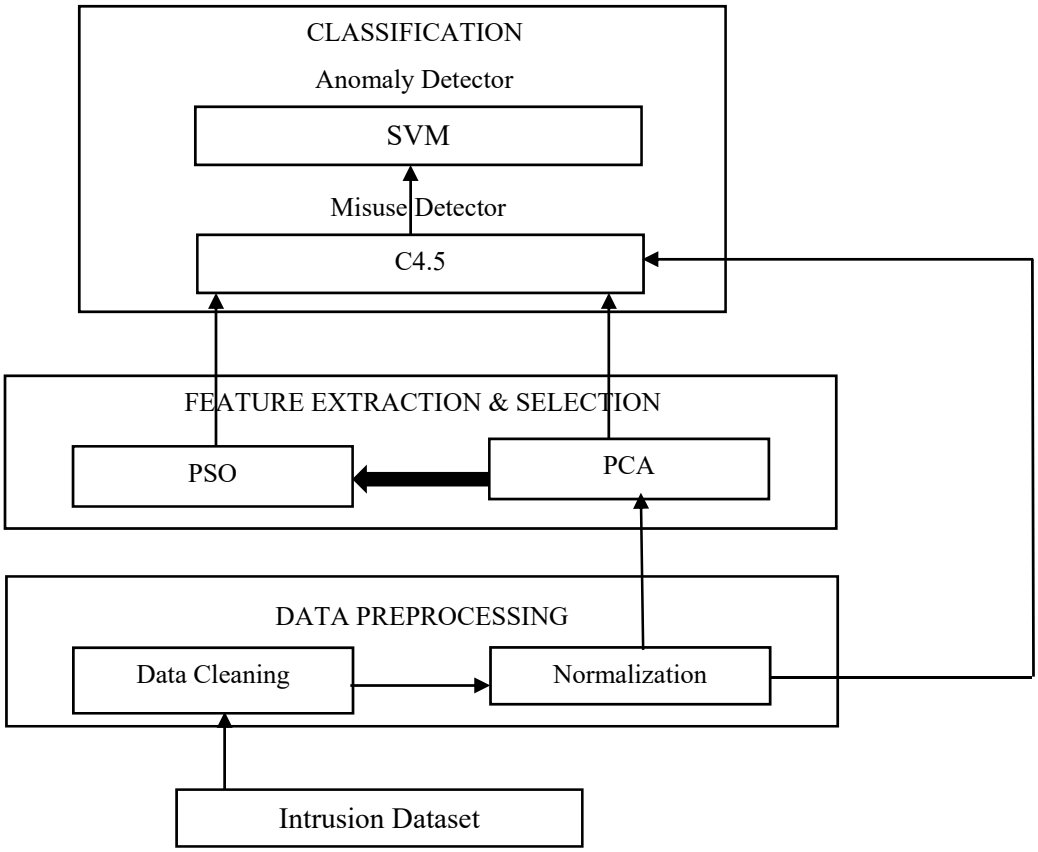

Fig. 1. System Architecture

3.3 Data Preprocessing

The intrusion raw dataset usually contains unprocessed and noisy data. For the raw data to be useful for classification purposes, it has to be cleaned and transformed into the form that could be useful for data mining

purposes [Zhang *et al*., (2003)]. Data preprocessing is the act of converting the original dataset into an understandable format. It aimed at reducing the vague in the original dataset to conform to the format that will enhance the prediction's accuracy. Also, it arranges the network data by grouping and handles the incomplete dataset. This stage involves data cleaning and data normalization.

3.3.1 Data Cleaning

The original dataset consisted of 41 attributes with one class label inclusive. In the raw dataset, all the symbolic features were first converted to a numeric value, the format appropriate for the dimensionality reduction algorithm, PCA. After that, the feature named 'NumOutboundCmds' and 'IsHostLogin' were completely filtered out due to their values being zero (0) and one (1) respectively, which would not have a significant contribution to the classifier's performance. The features were reduced from the initial 41 to 39 before applying data normalization. The dataset was converted and saved into database file Comma-Separated Values (CSV) format.

3.3.2 Data Normalization

Normalization of data is the process whereby the attribute in the data is arranged to enhance the entity's cohesion. It aimed at reducing and eliminating redundancy in the data model. The Min-Max Normalization technique was used to scale all the instances in respective features between the 0 and 1, and the formula is shown in Equation 1.

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Where X_{min} and X_{max}, represent the maximum and minimum value of a specific feature respectively, and X_n is a normalized output.

3.4 Feature Selection Stage

The reason for carrying out feature reduction and or selection is for the determination of the minimum number of features that would have high predictive power for classification purposes. The original feature set n is always greater than the selected subset feature p in the dataset. PCA algorithm was used for dimensionality reduction of the feature while PSO was used for the optimum selection of principal components from the PCA. Fig. 2 indicates the model for the optimized feature selection.

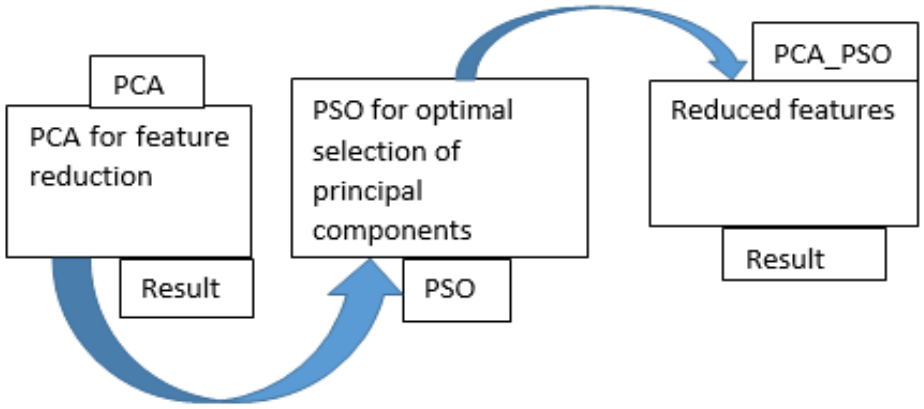

Fig.2. The model for the optimized feature selection

4. Result and Discussion

The proposed system is implemented and tested using on a Personal Computer (PC) with Intel(R) Core (TM) i3-2350M CPU @ 2.40GHz processor with 8 GB RAM running a 64-bit operating system. The developed IDS was built on the MATLAB IDE with the incorporation of the PCA for feature reduction, PSO for optimal selection of principal components from the PCA, C4.5 for misuse detector and SVM for anomaly detector. The NSL-KDD dataset was extracted from the Machine Learning Repository of UCI. The developed system followed a sequential procedure from data collection, data filtering, feature selection, network traffic classification, and detection. The system was simulated using MATLAB 2016A programming language employed to develop various classes and functions to link up the graphical user interface (GUI) for the interaction and responsiveness of the developed system. The developed systems made use of the various components and tools in Matlab to achieve the automation of this system. The implementation of the algorithm follows strictly from the

architecture developed in section 3. The automation of the created framework was coded and worked in a non-user interactive environment called the M-record. Different capacities and contents were created to make the created system a reality.

The training data consisted of twenty-five thousand, one hundred and ninety-two (25192) instances that were randomly selected from the NSL-KDD dataset. The selected dataset consisted of 11743 (46.6%) attacks and 13449 (53.4%) attacks. After the dataset was filtered, the attributes were reduced to 39. The system developed accommodated intrusion prediction by first reducing features with the PCA Algorithm and then selecting features with high predictive power using particle swamp optimisation technique. After selecting the relevant features, the dataset was passed to the misuse detector, C4.5 decision tree classifier and anomaly detector, SVM. The C4.5 was used to classify the traffic into the attack and normal; after that, the "assumed" normal was passed further to the second-level classifier (SVM) for further detection of zero-day (novel) attacks that the misuse detector might not have previously detected.

The system was developed to accommodate by first dimensionally reducing features with the PCA Algorithm and then selecting relevant features with high predictive power using PSO. After that, the results obtained were passed to the misuse detector, C4.5, to predict new instances. Later, the "assumed" normal was passed further to the second level classifier, SVM, for further detection of zero-day attacks that the misuse detector might not have detected. The PCA reduced the dataset from the original normalised dataset of 39 attributes to 24 features while PSO optimally selects relevant principal components from the PCA. The iteration was set to 1000 and the weight of 0.2. The weight is used to update the velocity function of the population. The PSO optimally select eleven (11) principal components from. The summary of the performance analysis of the system is presented in Table 1.

Table 1: Summary of the Analysis of the Feature Selection Methods for NSL-KDD Dataset

Feature Selection Method	**Selected Features**	**Accuracy Rate (%)**		**Detection Rate (%)**		**False Positive Rate (%)**		**F- measure**	
		C4.5	**C4.5_ SVM**	**C4.5**	**C4.5_ SVM**	**C4.5**	**C4.5_ SVM**	**C4.5**	**C4.5_ SVM**
Raw dataset	39	94.40	94.55	94.09	94.26	5.34	5.20	0.9394	0.941 1
PCA	24	96.26	96.36	95.95	96.05	3.45	3.37	0.9600	0.960 6
PCA_PSO	11	99.40	99.43	99.31	99.35	0.53	0.50	0.9936	0.993 8

Figure 3 presents the number of features selected as compared with the three selection techniques methods (raw data, PCA only and PCA_PSO) using a two-level classifier (C4.5 and C4.5_SVM). Out of the 39 features in the raw dataset, PCA dimensionally reduced the features into 24 features, while PCA_PSO optimally selected 11 features 24 features projected to the principal space by the PCA. Thus, the PCA_PSO provided the smallest relevant features which improve the performance of IDS by minimizing the prohibitive complexity of the two classifiers.

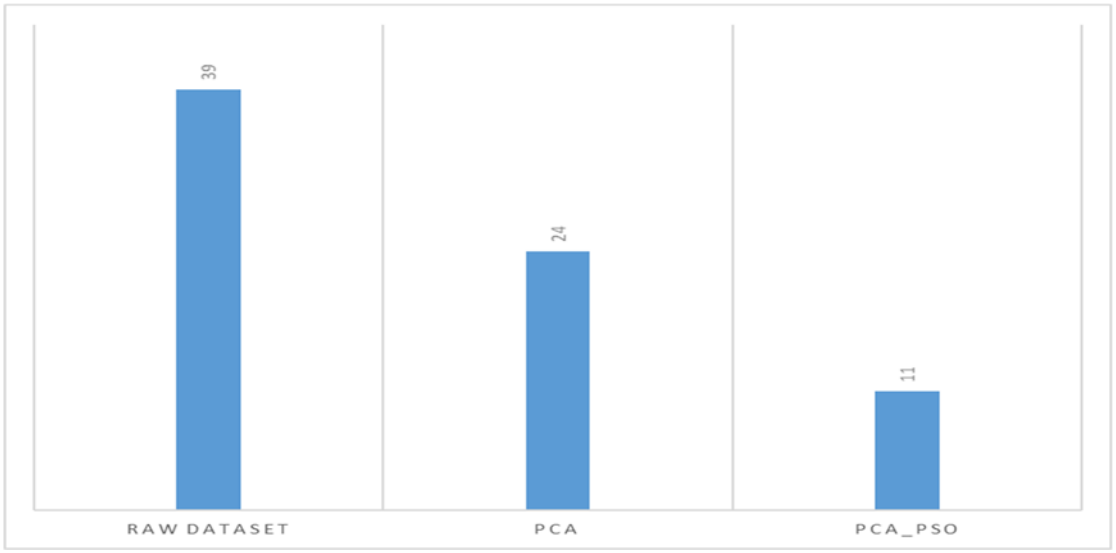

Fig. 3. Comparison of the selected features based on feature selection methods

Figure 4 presents the comparison of the accuracy rate of the various feature selection. The PCA_PSO feature selection method had the highest accuracy rate when compared with both raw dataset and PCA. The relevant features selected by the PCA_PSO had the highest predictive power for the classifiers. Using two-level classifier helped the IDS to have a very high accuracy rate. Therefore, the high value of accuracy rate of C4.5_SVM classifier implies that the developed IDS rightly classified as attack and normal.

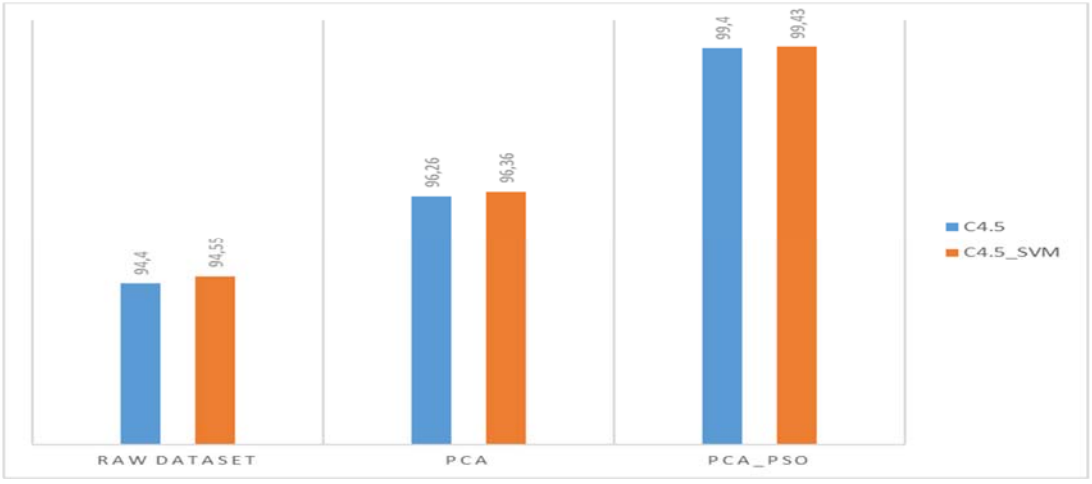

Fig. 4. Comparison of accuracy rate for feature selection methods using C4.5 and C4.5_SVM on NSL-KDD

Figure 5 compares the detection rate of all the three feature selection methods. The hybridised PCA_PSO feature selection method had the highest detection rate when compared with both raw dataset and PCA only feature selection methods. The relevant features selected by the PCA_PSO had the highest predictive power for the classifiers. Using a two-level classifier helped the IDS to have a very high detection rate. Therefore, the high value of detection rate of the C4.5_SVM classifier implies that the developed IDS accurately detected attack in network traffic.

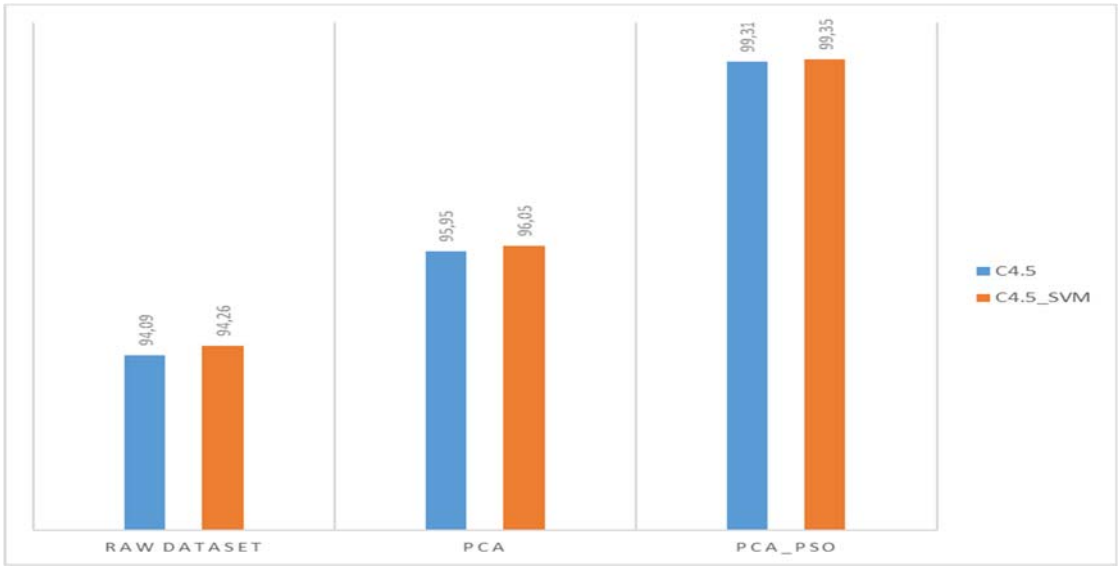

Fig. 5. Comparison of detection rate for feature selection methods using C4.5 and C4.5_SVM on NSL- KDD

As indicated in Fig. 6, the hybridised PCA_PSO feature selection method had the least false alarm rate when compared with both raw dataset and PCA only feature selection methods. The relevant features selected by the PCA_PSO had the highest predictive power for the classifiers. Using a two-level classifier helped the IDS to have a very low FAR. Therefore, the lower value FAR of C4.5_SVM classifier implies that the developed IDS minimally misclassified normal behaviour as an attack.

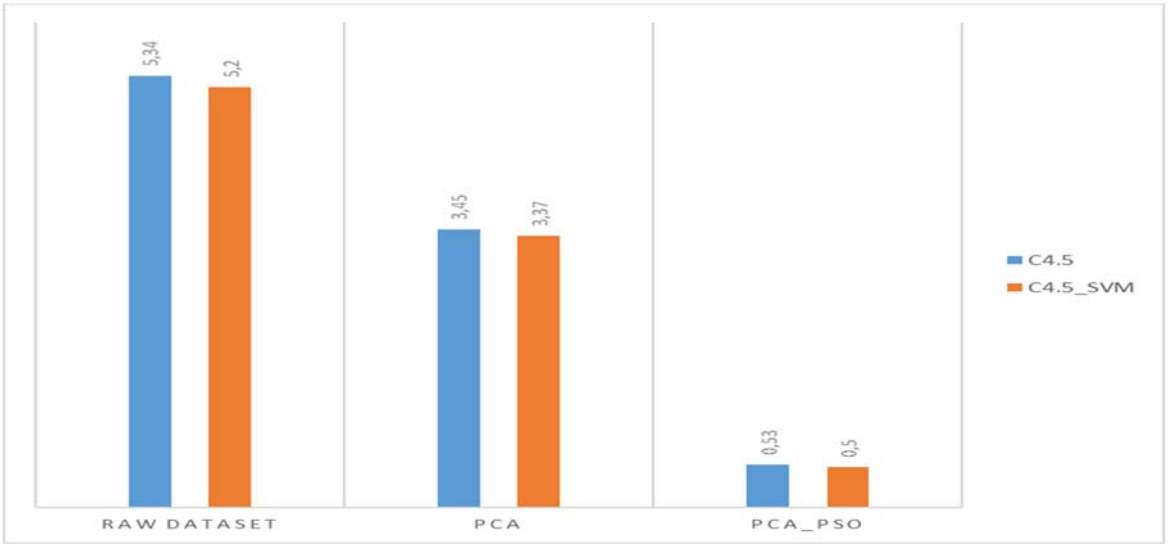

Fig. 6. Comparison of false alarm rate for feature selection methods using C4.5 and C4.5_SVM on NSL-KDD.

As indicated in Fig. 7, the hybridised PCA_PSO feature selection method had the highest precision, recall and f-measure value compared with both raw dataset and PCA only feature selection methods. The relevant features selected by the PCA_PSO had the highest predictive power for the classifiers. Using two-level classifier helped the IDS to have very high values for the precision, recall and f-measure. Therefore, the high precision and f-measure of the C4.5_SVM classifier implies that the developed IDS had a minimal false alarm.

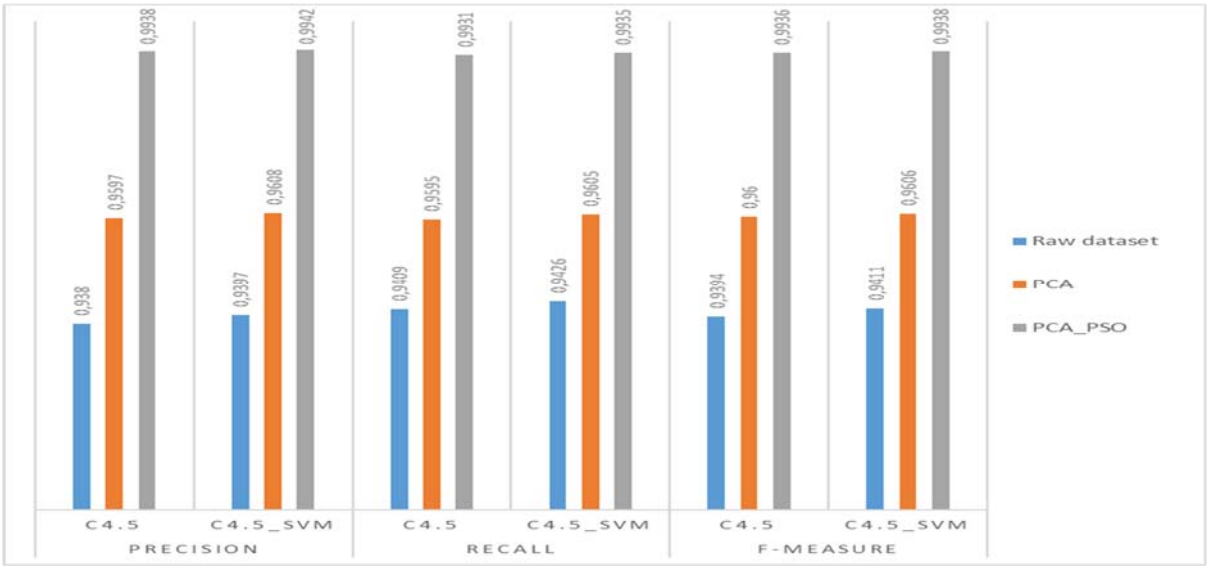

Fig. 7. Comparison of precision, recall and F-measure for C4.5 and C4.5_SVM on NSL-KDD

Table 2 presents the summary of the comparative analysis of the performance measures between NSL-KDD and KDD Cup'99.

Table 2: Summary of the Comparative Analysis of NSL-KDD and KDD Cup'99

Performance Metrics	NSL-KDD	KDD Cup'99
Accuracy Rate (%)	99.43	99.20
Detection Rate (%)	99.35	99.05
False Positive Rate (%)	0.50	1.06
Precision	0.9942	0.9874
Recall	0.9935	0.9950
F-Measure	0.9938	0.9912

Figure 8 presents the results of the performance metrics for both NSL-KDD and KDD Cup'99 datasets.

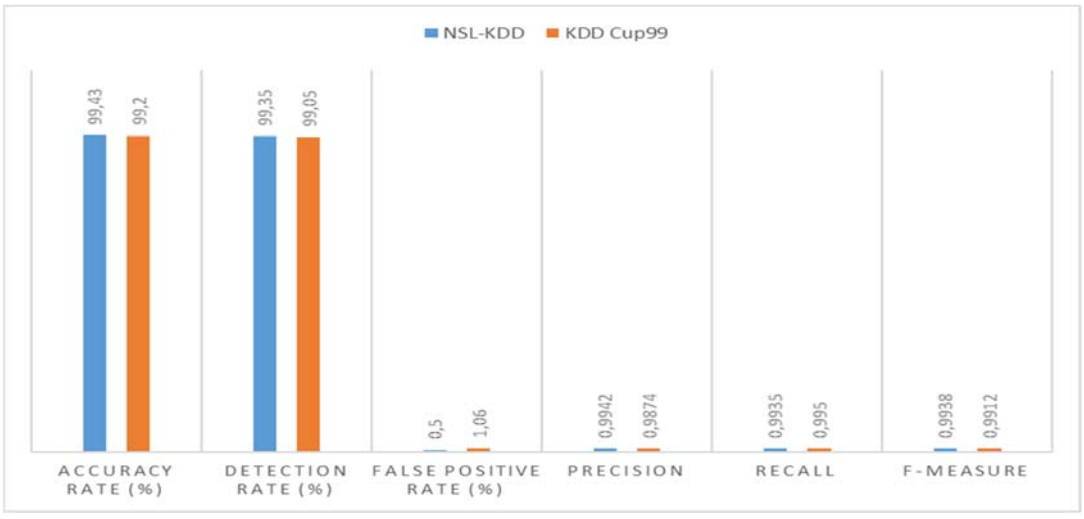

Fig. 8. Comparison of the performance of NSL KDD and KDD Cup'99

The performance analysis between NSL KDD and KDD Cup'99 showed that the accuracy rate for NSL-KDD and KDD was 99.43% and 99.20% respectively. The detection rate for NSL-KDD and KDD Cup'99 was 99.35% and 99.05% respectively. The false alarm rate of NSL-KDD and KDD Cup'99 was 0.50% and 1.06% respectively. The f-measure for NSL-KDD and KDD Cup'99 were 0.9938 and 0.9912 respectively. The results of the performance evaluation showed that NSL-KDD outperformed KDD Cup'99. This implies that NSL-KDD is a good synthesis intrusion dataset that can be used to evaluate the performance of IDS.

5. Conclusion

The paper presented a new hybrid feature selection and classification method. Based on the unprotected nature of some private and public networks, they are continuously exposed to security threats inside and outside of the networks. There called for a formidable measure to combat these security threats that can nip it in the bud. Various IDSs are available in the market but the problem is their performance. Hence, a feature reduction and selection method in IDS is proposed based on hybridising PCA and PSO to an optimal select subset of features projected to principal space based on component. The performance of the model was tested on both KDD Cup'99 and NSL KDD datasets which are considered the standard synthesis datasets for intrusion detection's evaluation. The experimental results showed that optimal feature subset selection performed effectively when measured in terms of accuracy, detection rate and false alarm rate and f-measure. This technique is easy to implement and performed efficiently in terms of the detection of network attacks.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] Agrawal, N.; Tapaswi, S. (2019): Defense mechanisms against DDoS attacks in a cloud computing environment: State-of-the-art and research challenges. IEEE Communications Surveys & Tutorials, **21**(4), pp. 3769-3795.
[2] Ahmad, I.; Abdullah, A.; Alghamdi, A.; Alnfajan, K.; Hussain, M. (2011): Intrusion detection using feature subset selection based on MLP. Scientific Research and Essays, **6**(34), 6804-6810.
[3] Al-hamami, A.; Alawneh, T. (2012): Developing a host intrusion prevention system using data mining. In International Conference on Advanced Computer Science Applications and Technologies (ACSAT), pp. 409-413.
[4] Anderson, J. P. (1980): Computer security threat monitoring and surveillance. Technical Report, James P. Anderson Company.
[5] Bovenzi, G.; Aceto, G.; Ciuonzo, D.; Persico, V.; Pescapé, A. (2020): December). A hierarchical hybrid intrusion detection approach in iot scenarios. In GLOBECOM 2020-2020 IEEE Global Communications Conference pp. 1-7. IEEE.
[6] Chandak, T.; Shukla, S.; Wadhvani, R. (2019): An analysis of "A feature reduced intrusion detection system using ANN classifier" by Akashdeep et al. expert systems with applications (2017). Expert Systems with Applications, **130**, pp 79-83.
[7] Chung, Y. Y.; Wahid, N. (2012): A hybrid network intrusion detection system using simplified swarm optimisation (SSO). Applied Soft Computing, **12**(9), pp. 3014–3022.
[8] Di Mauro, M.; Galatro, G.; Fortino, G.; Liotta, A. (2021): Supervised feature selection techniques in network intrusion detection: A critical review. Engineering Applications of Artificial Intelligence, **101**, pp. 104-216.
[9] Ektefa, M.; Memar, S.; Sidi, F.; Affendey, L. S. (2010): Intrusion detection using data mining techniques. In International Conference on Information Retrieval and Knowledge Management, pp. 200-203.
[10] Elngar, A. E.; Mohammed, D. A.; Ghaleb, F. F. M. (2013): A real-time anomaly network intrusion detection system with high accuracy. An International Journal of Information Sciences Letters, **2**, pp. 49-56.
[11] Eskandari, M.; Janjua, Z. H.; Vecchio, M.; Antonelli, F. (2020): Passban IDS: An intelligent anomaly-based intrusion detection system for IoT edge devices. IEEE Internet of Things Journal, **7**(8), pp 6882-6897.
[12] Fernandes, G.; Rodrigues, J. J.; Carvalho, L. F.; Al-Muhtadi, J. F.; Proença, M. L. (2019): A comprehensive survey on network anomaly detection. Telecommunication Systems, **70**(3), pp. 447-489.
[13] Folorunsho, O.; Jimoh, R. G. (2017): A framework for implementing hybrid intrusion prevention model based on soft-computing techniques. In Proceeding of the 11th International Multi-Conference on ICT Applications, AICTTRA, pp. 139 -142.
[14] Gu, J.; Wang, L.; Wang, H.; Wang, S. (2019): A novel approach to intrusion detection using SVM ensemble with feature augmentation. Computers & Security, **86**, pp. 53-62.
[15] Hider, W.; Hu, J.; Slay, J., Turnbull, B. P.; Xie, Y. (2017): Generating realistic intrusion detection system data based on fuzzy qualitative modeling. Journal of Network and Computer Applications, **87**, pp. 185-192.
[16] Ingre, B.; Yadav, A.; Soni, A. K. (2017, March): Decision tree based intrusion detection system for NSL-KDD dataset. In International conference on information and communication technology for intelligent systems pp. 207-218. Springer, Cham.
[17] Juumi, K.; Wookey L.; Justin J. S.; Soo-Bok, L. (2017): Optimised combinatorial clustering for stochastic processes. Cluster Computing. **23**(2), pp. 5–16.
[18] Khraisat, A.; Gondal, I.; Vamplew, P.; Kamruzzaman, J.; Alazab, A. (2020): Hybrid intrusion detection system based on the stacking ensemble of c5 decision tree classifier and one class support vector machine. Electronics, **9**(1), pp. 1-18.
[19] Krishnaveni, S., Vigneshwar, P.; Kishore, S.; Jothi, B.; Sivamohan, S. (2020): Anomaly-based intrusion detection system using support vector machine. In Artificial Intelligence and Evolutionary Computations in Engineering Systems, pp. 723-731. Springer, Singapore.
[20] Kumari, R.; Sharma, K. (2018): Cross-layer based intrusion detection and prevention for network. In Handbook of Research on Network Forensics and Analysis Techniques pp. 38-56. IGI Global.
[21] Liu, H.; Lang, B. (2019): Machine learning and deep learning methods for intrusion detection systems: A survey. applied sciences, **9**(20), pp. 1-28.

[22] Maglaras, L. A.; Jiang, J. (2014): Intrusion detection in SCADA systems using machine learning techniques. In Proceedings of Science and Information Conference of IEEE, 626-631.
[23] Manzoor, I.; Kumar, N. (2017): A feature reduced intrusion detection system using ANN classifier. Expert Systems with Applications, **88**, pp. 249-257.
[24] Mirsky, Y.; Doitshman, T.; Elovici, Y.; Shabtai, A. (2018): Kitsune: an ensemble of autoencoders for online network intrusion detection. arXiv preprint arXiv:1802.09089.
[25] Mohammed, R.G.; Awadelkarimi, A.M. (2011): Design and implementation of a data mining network intrusion detection scheme. Asian Journal of Information Technology, **10**(4), pp. 36-141.
[26] Mohammadi, S.; Amiri, F. (2019): An efficient hybrid self-learning intrusion detection system based on neural networks. International Journal of Computational Intelligence and Applications, 18(01), 1950001.
[27] Moustafa, N.; Slay, J.; Creech, G. (2017): Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks. IEEE Transactions on Big Data, **5**(4), pp. 481-494.
[28] Moustafa, N.; Turnbull, B.; Choo, K. K. R. (2018): An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of the internet of things. IEEE Internet of Things Journal, **6**(3), pp. 4815-4830.
[29] Nadiammai, G. V.; Hemalatha, N. (2013): Handling intrusion detection system using snort statistical algorithm and semi-supervised approach. Research Journal of Applied Sciences, Engineering and Technology, **6**(16), pp. 2914-2922.
[30] Nimbalkar, P.; Kshirsagar, D. (2021): Feature selection for intrusion detection system in Internet-of-Things (IoT). ICT Express, **7**(2), pp. 177-181.
[31] Ogonji, M. M.; Okeyo, G.; Wafula, J. M. (2020): A survey on privacy and security of Internet of Things. Computer Science Review, **38**, pp. 100-112.
[32] Pradhan, M.; Nayak, C. K.; Pradhan, S. K. (2020): Intrusion detection system (IDS) and their types. In Securing the Internet of Things: Concepts, Methodologies, Tools, and Applications (pp. 481-497). IGI Global.
[33] Ramakrishnan, S.: Devaraju, S. (2017): Attack's Feature Selection-Based Network Intrusion Detection System Using Fuzzy Control Language. International journal of fuzzy systems, **19**(2), pp. 316-328.
[34] Rathore, M. M.; Saeed, F.; Rehman, A.; Paul, A.; Daniel, A. (2018, February): Intrusion detection using decision tree model in high-speed environment. In 2018 International Conference on Soft-computing and Network Security (ICSNS). pp. 1-4. IEEE.
[35] Salo, F.; Nassif, A. B.; Essex, A. (2019): Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection. Computer Networks, **158**, pp. 164-175.
[36] Sheen, S.; Rajesh, R. (2008): Network intrusion detection using feature selection and Decision tree classifier. In Proceedings of 10th International Conference of TENCON IEEE Region, pp. 1-4.
[37] Shenfield, A.; Day, D.; Ayesh, A. (2018): Intelligent intrusion detection systems using artificial neural networks. Science Direct (ICT Express), **4**, pp. 95-99.
[38] Soheily-Khah, S.; Marteau, P. F.; Béchet, N. (2018): Intrusion detection in network systems through hybrid supervised and unsupervised mining process: A detailed case study on the ISCX benchmark dataset. In the Proceedings of 1st International Conference on Data Intelligence and Security, pp. 219-225.
[39] Stehlik, M.; Matyas, V.; Stetsko, A. (2017): Attack detection using evolutionary computation. In Computational Intelligence in Wireless Sensor Networks pp. 99-129. Springer, Cham.
[40] Subba, B.; Biswas, S.; Karmakar, S. (2016): A neural network-based system for intrusion detection and attack classification. In Proceeding of 22nd National Conference on Communication (NCC), 1–6.
[41] Tesfashun, A.; Bhaskari, D.L. (2017): Effective hybrid intrusion detection system: A layered approach. International Journal of Computer Network and Information Security, **3**, pp. 35-41.
[42] Xue, Y.; Jia, W.; Zhao, X.; Pang, W. (2018): An evolutionary computation based feature selection method for intrusion detection. Security and Communication Networks, 2018.
[43] Zarpelao, B. B.; Miani, R. S.; Kawakani, C. T.; Alvarenga, S. C. (2017): A survey of intrusion detection in the Internet of Things. Journal of Network and Computer Applications, **84**, pp. 25-37.
[44] Zhang, S.; Zhang, C.; Yang, Q. (2003): Data preparation for data mining. Applied artificial intelligence, 17(5-6), 375-381.

Authors Profile

FOLORUNSHO, **O.** is currently a Postdoctoral Research Fellow at the Unit for Data Science and Computing, North-West University, Potchefstroom, South Africa and lecturer at the Department of Computer Science, Federal University Oye Ekiti, Nigeria. He obtained his PhD at the University of Ilorin, Nigeria, Master of Science (M.Sc) degree of the University of Ibadan, Nigeria, Postgraduate Degree in Education (PGDE) of Usmanu Danfodiyo University Sokoto, Nigeria and a Bachelor of Technology (B.Tech Hons) degree of the Federal University of Technology, Minna, Nigeria. He is a member of the Computer Professionals Registration Council of Nigeria and Nigerian Computer Society (NCS). His research interests include Information Security, Data Mining, and Artificial Intelligence.

ADEGBOLA I. A. is currently the Head of Department, Computer Science, Oyo State College of Education Lanlate, Nigeria. He holds BSc. Computer Science (Ilorin) Second Class Honours Upper Division, MSc. Computer Science (Ibadan), Ph.D Computer Science (Ilorin) and Professional Diploma in Education F.C.E (Special). He is a member of Nigeria Computer Society (MNCS), member, Teacher Registration Council of Nigeria.

JIMOH, R. G. is currently a full Professor in the Department of Computer Science, Faculty of Communication and Information Sciences, University of Ilorin, Ilorin. He holds National Diploma Computer Science (Offa) (Upper Credit and as best graduating student), B.Sc. Computer Science (Ilorin) (Second Class Honours (Upper Division), M.Sc. Computer Science (Ibadan), Ph.D Computer and Information Technology (UUM, Malaysia). He is a member of Computer Professionals Registration Council of Nigeria (MCPN), Member, Nigeria Computer Society (MNCS). He is currently the North-central Coordinator and Representative to the Council of Computer Professional Registration Council of Nigeria.