International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 4 Apr 2022, pp: 868-877 www.ijaem.net ISSN: 2395-5252

Machine Learning Intrusion Detection System: Opportunity and Challenges

Ohuabunwa Augustine Ebere[1*], Okeke Obinna[2], Icheke Led[2], Okwuelu Nnaemeka[2], Kadiri Ramotu Ochuwa[1], Mekudi Sunday[2]

[1]Information Communication Technology Unit, Electronics Development Institute, Awka, Nigeria
[2]Research and Development Unit, Electronics Development Institute, Awka, Nigeria
Corresponding Author: Ohuabunwa Augustine Ebere

Submitted: 05-04-2022 Revised: 16-04-2022 Accepted: 19-04-2022

ABSTRACT: Presently, machine learning (ML) is enjoying an unparalleled surge in applications that solve network security problems and facilitate automation in diverse areas. This is attributed to the sudden explosion in the volume of available data, remarkable improvement in ML techniques, and innovation in computing and networking technology. It is employed in providing solution to both mundane and complex problems such as intrusion detection in information system. Machine Learning Intrusion detection systems (IDSs) are among the most promising tools for securing data and networks. In the last century, many classification algorithms such as neural network (NN), decision tree (DT), Naive Bayes (NB) and support vector machine (SVM) have been applied in IDS. However, each of these algorithms is associated with the challenges of tuning the parameters of the classifiers correctly and choosing the best subset features for detection. To overcome these challenges, optimization algorithms like genetic algorithm (GA), ant colony algorithm, particle swarm optimization (PSO) algorithm and other algorithms are used together with classifiers to make them more efficient. Through a systematic literature review of research on machine learning classifiers and optimization algorithm used in Intrusion Detection System; this paper aim at dealing with the issues associated with machine learning intrusion detection system and present future perspectives for further research in the field of machine learning IDS by exploring the measures suggested in existing studies.

KEYWORDS Algorithm, Intrusion, Detection, Security, Model, Machine learning, Anomaly, attack, Optimization and Classifier

I. INTRODUCTION

With the wide spread of social networking site and the increasing rate of business transactions over the Internet, network security has become a critical issue [1]. Despite the commitment to building a secure system, Information and communication system security vulnerabilities surface almost daily. Thus, there is the need for network managers to take proactive measures to defend and ensure quick recovery of the system in event of attack. In small cyber-community, attack prevention can be possible. However, the advancement in hacking technologies and availability of numerous free hacking tools has made cyber attack prevention almost an impossible task. Given sufficient time and assets, any network can be compromised. The most frequent risk to a network's security is an intrusion like brute force and denial of service from external attackers or an infiltration from within a network; this compromise the availability, integrity and confidentiality of the system. With the shifting patterns in network behaviour, it is essential to switch to a more dynamic approach to identify and prevent such intrusions.

Over the years, intrusion detection system has been developed using various techniques such as statistics, provenance graph, specification, machine learning etc. Generally, there are two distinctive methodologies for modelling intrusion detection namely, anomaly and misuse. Anomaly detection assumes that normal user profile is absolutely observable and completely different from intrusive. The model is developed using the normal user behaviour and any user profile which differs from the established one is flagged as an intrusion [2]. It has the ability to detect novel attack with an inherent problem of high false alarms since evolving normal behaviours is seen as an attack. The misuse detection use signature of well-known threats to search for a match in the monitored data; if there is a match, an intrusion is reported. It has minimal false alarm but detect only the attacks that has its signature stored in the attack database.

To overcome the challenges of both the anomaly and misused detection, hybrid detection which mostly employs machine learning technique was introduced. Machine Learning has the capacity to identify and exploit hidden patterns in "training" data. The patterns learnt are use to analyse unknown data, such that it can be grouped together or mapped to the known groups. It is classified as supervised or unsupervised. When using supervised learning techniques, the algorithm is trained using a labelled datasets to derive a function capable of mapping instances to classes. In the case of unsupervised technique; instead of training, the inference engine is presented with data item which is classified into a set of groups taking advantage of the intra-group and inter-group relationship.

II. MACHINE LEARNING ALGORITHMS OVERVIEW.

Machine learning is a field of study that enable computers to learn and carryout a given task without the need of first being programmed [3]. Network security applies ML in decisions making to detect the presence of malicious packets. This section will focus on ML algorithms which is utilize in building IDS to explore the capabilities and limitations.

2.1. K-Nearest Neighbour

The K-NN algorithm is one of the most fundamental classification approaches. To classify a given dataset, the algorithm looks at the k-nearest neighbour and vote for them on the basis of the calculated measurement (Euclidean distance between a given test sample and a specified training sample of data) [4].The votes given to the majority value of the nearest neighbour determines the group to classify the dataset. Using k-nearest neighbour, it is easy to implement and learn complex functions effortlessly. However, it suffers from high dimensionality and over-fitting.

2.2. Artificial Neural Networks (ANN)

ANN models are built to mimic the functionality of the biological neural network [5]; they usually contain many layers in which the output of one layer becomes the input for the succeeding layer. During classification, the last output of the output layer generates the final classification category. NNs guarantees better pattern recognition than other algorithms. Its self-organizing maps do not require label at input; though, it needs more time for training and has low detection precision for low frequent attacks.

2.3. Decision Trees (DT)

Decision tree as a machine learning algorithm has three vital components: a decision node which specifies the test attribute, an edge that characterize the possible attribute values, and a leaf that represent the class of the object [6]. Based on the DT model, new instances are classified beginning from the root and moving to the branches based on the associated value of the considered object. An instance is said to be classified on reaching a leaf [6]. DT achieves high accuracy with better false positive than most other machine learning algorithm.

2.4. Bayesian

The Bayesian classifier (BC) relies on the idea that a (natural) class is expected to predict the feature values for its members [6]. It is the best algorithm to use when training data is scarce. Bayesian algorithm has the advantage of reduce time complexity. It is simple and guarantees high accuracy.

2.5. Support Vector Machine (SVM)

SVM is a supervised machine learning technique basically used for classification and prediction [7, 8]. To classify data, the algorithm maps an input vector into a predefined high-dimensional feature space through some nonlinear mapping [9]. The feature space is constructed by means of an optimal separating hyper-plane. SVM is effective in high-dimensional spaces. It separates hyper-planes into two classes: +1 and −1, with the positive value representing normal data and the negative anomalous data.

2.6. Challenges of Machine Learning IDS

Challenge 1. The need to obtain a minimal feature (input variable or attribute) set that is comprehensive enough to accurately separate normal data from intrusive data. According to [10], dimensionality and size of dataset affect the training complexity of machine learning algorithm; hence feature selection mechanism need to be apply [11, 12]. Feature selection reduces the dimension of processed data which in turn reduces the complexity and time needed to process data for intrusion detection.

Challenge 2. Most machine learning algorithms require specific parameters that need to be controlled and this make the operation to take long time [13] as the various frameworks require specific control parameters. For instance, GA has mutation and crossover parameter as the selection operator whereas PSO depends on inertia weight, cognitive and social parameters. The performance of these algorithms depends on how well the specific parameters are tuned.

III. SINGLE MACHINE LEARNING IDS TECHNIQUE

A lot of research has been carried out on IDS which employ single machine learning algorithm. Chan et al. [14] presented FAR and FAP IDPS to mitigate web service attacks, particularly attack against software-as-a-service cloud computing. The IDPs/IPSs which were deployed over a public cloud platform used 336 fuzzy association patterns in addition to 20 fuzzy association rules. The system performance was evaluated based on the transaction time in various operational scenarios. The two fuzzy-based detection systems detected and prevented famous web attacks with detection accuracy close to 100% and false alarm rate which is less than 1%. Also, Wang et al. [15] proposed a botnet detection technique based on behaviour analysis using fuzzy pattern identification scheme. By adjusting accurately the relationship functions employed in the fuzzy pattern, the authors successfully improved the detection accuracy. The performance of the proposed scheme was above all the baselines parameters used in evaluating the result. The false positive alarm rate is below 3.5% and it is capable of detecting about 95% of bots.

Watson et al. [16] developed an anomaly-based intruder detection technique which employs the SVM algorithm to identify malware in cloud's hypervisor level. The authors extracted the features and generate feature set using end system and network-level data. Their scheme achieved 90% anomalies detection accuracy using system-based features but the results are less accurate with network-level features. However, this result is an indication that SVM method can be used for malware detection with minimum time cost.

Protection scheme against DDoS attacks was introduced by Iyengar et al. [17] using fuzzy logic. The authors examined the various types of DDoS attacks and provided solution to secure the cloud environment through a fuzzy system IDS which monitors the network input traffic to detect DDoS attacks. At the first working stage of the system, the traffic types were specified via the fuzzy system design rule. Then, there is traffic analysis and evaluation. On detecting anomaly, the system triggers alarm and send a request to the routers to block the malicious packet. This method decrease storage functionality and cost of data transmission.

Wang et al. [18] proposed a hybrid IDS for mobile malware detection. They utilized the signature-based technique for the detection of known attack while the anomaly-based scheme using linear SVM classifier was employed for zero-day attack detection. Initially, the SVM-based detection checks any new application and classifies it as normal or abnormal. If it is abnormal, the signature detection will distinguish the type of abnormality. Using classification rate, the results obtained were evaluated for each malware family. When, the result is compared with two variants of the Support Vector Classifier (SVC), it was observed that the new method has the highest classification rate. The false-negative rate and true-positive rate are 1.16% and 98.94% respectively. However, the performance of this system is adversely affected by the presence of native codes and applications incorporating HTML5.

Khorshed et al. [19] studied numerous security issues in cloud computing and presented a proactive approach to identify the threats. It obtains its information from the threat sample and alerts the administrator or the user about the threat. Authors used SVM scheme for the detection of cloud attacks and compared the result with other machine learning methods like DT [20], PART [21], multilayer perceptron [22], and Naïve Bayes [23]. Their work revealed that when compared in terms of accuracy and processing time, SVM exhibits the best performance for the attack detection task against the other machine learning techniques. Also, the authors examined the SVM variants: polynomial, normalized polynomial, and RBF kernel types. It was observed that the SVMs with the polynomial kernel provide the highest accuracy, whereas the RBF-based SVM performance was the lowest. Furthermore, increasing the polynomial degree tends to decrease the accuracy and increase the processing time.

Pitropakis et al. [24] introduced a network attack detection technique that utilizes the Genetic Algorithm presented in [25]. The authors showed that attacks and malicious activities could be detected in a cloud setting by examining the system calls produced during the various steps of the attack and contrasting the system calls with other execution of the same attack, also with the usual system state when the attack took place.

Study on the detection of Virtual Machine-to-hypervisor attack was carried out by Nezarat et al. [26]. The authors described a Game Theory (GT) based IDS scheme in which many agents distinguished the attack and its origin leveraging on the Nash equilibrium concept. They also evaluated the performance of the GT-based method using the attack detection rate and compared it with the result obtained applying other machine learning approach such as: parallel neutral networks, genetic algorithms and neural networks, genetic algorithm and fuzzy logic, and a model for service-oriented architectures. The result showed that the GT-based algorithm reduced both the systems overhead and the number of false alerts drastically while at the same time increasing the attack detection rate up to 86%.

The work by Osanaiye et al. [27] is aimed at increasing classification precision and reducing the computational complexity of machine learning IDS by introducing an ensemble-based method (EMFFS). It focused at eliminating the redundant features during the pre-processing stage in order to speed up data classification via a decision tree. By using this approach in differentiating attacks from normal traffic in cloud computing, the authors achieved greater efficient learning time, higher detection rate and lesser complexity. Also, the decision tree trained using the feature set derived from the EMFFS method have the highest accuracy in DDoS attack detection when compared with linear correlation-based feature selection (CFS)[28], gradual feature removal [29] and CFS consistency-based subset evaluator [30]. In addition, the authors revealed that EMFFS can reduce the number of features from 41 to 13 and hence reduce the complexity of the classification task in comparison to other classification technique.

Kumar et al. [31] proposed a cloud IDS which centered on clustering using automata for healthcare vehicular cloud computing. By using aggregate relative velocity and connectivity degree, the learning automata formed the leadership of clusters. Also data is secured by utilizing the Hash based Message Authentication Code (HMAC) algorithm to validate messages. This method easily adapts to variation in the position of nodes in the network, produce lower false-positive rate and detect approximately 93% of malicious activities in the network.

Huang et al. [32] presented an anomaly detection technique employing Local Outlier Factor (LOF) and dimension reasoning rules. Using this approach the rate of anomaly detection was increased by 98%, and at the same time the false alert rate was decreased to 16.9% base on the classic LOF used as a baseline. The approach is more effective for virtual machine management since it detect anomalies via virtual machine's performance profile and give report base on the origin.

Sharma et al. [33] introduced an IDS technique base on artificial bee colony for the detection of DDoS attack. The main objective of this work is to demonstrate the efficiency of artificial bee colony method in the detection of DDoS attack. The results were assessed based on accuracy of attack detection and compared with a quantum behaved PSO (QPSO) baseline. The authors achieved approximately 8% increase in accuracy compared with PSO baseline.

Muthukumar and Kumar [34] employed artificial intelligent technique in developing IDS for a private cloud. The IDS components were firstly trained and then tested ascertain that the training stage was completed successfully. Finally, the system was updated accordingly. The results showed that the new technique could improve the performance of IDS used in private cloud computing environment with regards to time and space complexity.

Chiba et al. [35] developed a novel NIDS scheme integrating signature-based and anomaly-based detection, an enhanced Back-Propagation Neural Network (NN) and Snort IDS [36, 37]. Firstly, Snort inspects received packets. If an intrusion is discovered, it sends an alert; otherwise the packet is transmitted to the anomaly detection phase. At this point, a back-propagation NN verifies the type of packet to ascertain if it is normal or abnormal. The alerts produced by the system are sent to a database which is used to discover the intrusions. This approach is capable of increasing the detection accuracy while minimizing the rates of false negatives and false positives. In addition, it guarantees an appropriate cost for computations.

Ghosh et al. [38] presented a technique to minimize the size of the dataset by a combination of Nearest Neighbour reduction and Penalty-Reward based instance selection. To show the effectiveness of this method, the authors compared the new scheme with NN, Ada-Boost, and Random Forest classifiers on the original and reduced dataset. The result showed that the data-optimization technique developed not only reduced the training time (due to the data-dimensionality reduction) but also improved classification accuracy for the designed IDS

Chonka et al. [39] introduced Cloud Protection System which is a defence mechanism based on a back-propagation NN for Cloud Computing. This scheme dictates 98-99% of the XML- and HTTP-DoS attacks within 10-135ms in addition to identifying the attack's origin.

In [40], the researchers provided a detection scheme for cloud and grid computing which makes use of anomaly- and misuse-based detection to verify attacks. The authors employed an artificial neural network (ANN) for their anomaly-based detection whereas the communication and log systems' data were analysed for misuse-based detection. Their experiment result showed that the false positive is lower than the false-negative rate. Besides, the proposed model led to low data volume and complexity requirements, and gives satisfactory performance for real-time implementation.

Xiong et al. [41] proposed an anomaly detection approach which examines the dynamic characteristics of the network traffic using catastrophe theory [42] and synergetic neural network [43] for a cloud computing environment. The authors applied these two techniques separately, showing that they have the capacity to detect anomalous traffic in any given network. Particularly, catastrophe theory can identify unexpected variation in the network traffic and then discover anomalies associated with the deviation of the state of the network traffic from the usual one. The synergetic neural network is a pattern recognition process which performs anomaly detection by matching the testing data with the training data. They proved that both schemes improved the rate of false alerts as well as detection probability over compared baseline.

3.1 Optimization Algorithms

3.1.1 Genetic Algorithm

Genetic algorithm as an optimization method is based on evolutionary computation and depends on the survival of the fittest concept. It relies on two operators: mutation and crossover to achieve optimality [44]. Senthilnayaki et al. [45], proposed a machine learning IDS model based on SVM and GA algorithm. The researchers used GA to select 10 optimal features out of the 41 features in KDDCUP99 dataset. The IDS was evaluated and found to performance better in terms of accuracy and classification time than when the entire 41 features were used.

3.1.2 Particle swarm optimization algorithm (PSO)

Particle Swarm optimization algorithm was introduced in 1995 by Eberhart and Kennedy [46] based on insight from social behaviours, such as flocking of birds and schooling of fish [6]. PSO is similar to evolutionary computation but lacks the evolution operators. Several studies have attempted the combination of SVM and PSO to enhance IDS performance. Wang et al. [47] suggested a novel IDS using PSO algorithm and SVM. In their work, the PSO algorithm was employed to optimize SVM parameters. The result obtained after the system was evaluated showed a remarkable improvement in the performance of the IDS; there is reduction in the classification time in addition to increase in accuracy and detection rate [47].

3.1.3 Ant colony algorithm (ACO)

Dorigo et al. [48] proposed ant colony algorithm for solving difficult optimization problems. The algorithm was developed from inspiration on the high level of organization in ant colonies. Ants have the ability to find the shortest path to their colonies by depositing pheromone down their path to lead the subsequent ants toward their food source or colony. The pheromone deposited by the ants down their path disappear with time; therefore, paths with less pheromone will be less popular while the path frequently travelled by ants will continue to have higher pheromone concentration and many ants will keep choosing it as the shortest route to their colony. Gao et al. [49] presented IDS model that combined ACO (for Feature Selection) and SVM (for Intrusion Detection). Using this scheme, the intrusion features are represented as graph nodes whereas the edges denote the addition of new features. By ensuring that some SVM classifiers are not trained; the method minimizes the number of features and improved SVM performance.

3.2 Hybrid machine learning IDS technique

Several works have been done with regards to Hybrid machine learning IDS technique. Ganeshkumar and Pandeeswari [50] introduced an adaptive neuro-fuzzy inference system (ANFIS) which is a hybrid scheme for intrusion detection based on fuzzy systems and neural network. This is a Hypervisor Introspection and can detect network-based intrusion as well as host-based intrusion without being deployed directly in the VMs. The authors evaluated the system performance using DARPA's KDD Cup dataset assuming five types of attacks. They adopted precision, recall and F-measure values as the test parameter and compared the result obtained with Naïve Bayes, NBRF and ANN scheme. ANFIS gave a recall which is comparable with other methods; however, it has higher precision and F-measure ($\simeq$ 97). ANFIS is designed for Big Data applications.

Pandeeswari and Kumar [51] presented another hybrid technique for the hypervisor layer of cloud systems by combining both artificial neural network and Fuzzy C-Means clustering algorithm (FCM-ANN). The model has three phases: fuzzy clustering module, ANN module, and fuzzy aggregation module. Initially, the system cluster data into small groups to boost the ANN's learning ability. Next is the training of ANN components using the value of the distinct clusters. Finally, the outcome of the ANN is incorporated by the aggregation module. When using this model, there is no need to capture the attack patterns manually. The authors also used DARPA's KDD Cup dataset in carrying out an experiment to evaluate the proposed system and compare the result obtained with that of similar models. The result indicated that FCM-ANN have the highest precision (about 65%), recall (≃90%) and F-measure (≃75%), outperforming both Naives Bayes classifier and standard ANN.

The authors in [52] developed a new IDS model by integrating Particle Swarm Optimization and Bayesian networks. Their experiment was conducted using KDD CUP 99 dataset and the scheme gave better result in terms of false-positive rate, detection time and detection rate when compared with similar IDS.

Raja and Ramaiah [53] suggested an intrusion detection system with the combination of Genetic Algorithm (GA) and Fuzzy NN (ANN-GA). The authors applied GA to overcome the rate of detection problem associated with Fuzzy NN to differentiate between users-to-root and remote-to-local attacks. Initially, clustering is applied using k-means algorithm [54] to achieve high precision. Next, a GA scheme [55] is employed to extract and optimize the fuzzy rules. Finally, the Fuzzy NN carries out the refinement of parameters. After this sequence of activities, the authors generated a rule base for intrusion detection. The result obtained when the hybrid ANN-GA is compared with an FNN [56], the two types of ANN proposed in [57] and [58], and two variant of the GA according to [59] and [60] using a standard IDS benchmark data showed that the proposed ANN-GA scheme has the best average detection accuracy.

Ghosh et al. [61] proposed a hybrid IDS integrating multi-threaded Host-based Intrusion Detection System (HIDS) and Network-based Intrusion Detection System (NIDS). The packet analyser utilizes a hybrid NN with K-Nearest Neighbour approach (KNN-NN) to examine the network traffic. The system takes into consideration both misuse and anomaly detection. Firstly, the captured network packets are sent to the KNN-NN classifier which group packets into normal or abnormal. Then, an ANN examines just the abnormal packets to establish the type of attack. Also, the authors were able to detect hypervisor attacks using HIDS. The combination of HIDS and NIDS give rise to a reliable and secure system which is faster and more efficient than rivals.

The authors in [62] developed an anomaly detection technique based on a clustering algorithm to detect abnormal virtual machines (VMs) (for instance, VMs corrupted with malicious software which is attacking others). The authors came up with a feature extraction algorithm which employs Locality Preserving Projections [63] and Principal Components Analysis [64] to decrease the data dimensionality. The anomaly detection utilizes a distance-based clustering algorithm supplied with the extracted features. The results of the experiment performed showed that the proposed method has higher efficiency in terms of precision, false alarm rate, recall and runtime.

Idhammad et al. [65] used data mining method to design a Distributed Intrusion Detection System (DIDS) for cloud computing. The system comprises of five components used for the collection of network traffic, data pre-processing, anomalies detection, synchronization of identified malicious data, and classification of attacks. Every router on the network boundary gathers traffic data and forwards them to the pre-processing unit which uses a time-based sliding window algorithm in processing and normalizing the data. Next, the anomaly detection unit categorizes the network traffic as normal or abnormal using the Naïve Bayes algorithm. After the initial anomaly detection per time window, the malicious traffic coming from each of the routers is subsequently synchronized to a centralized storage server. Finally, the attack type is detected by utilizing a Random Forest classifier. The performance of the system was evaluated using CIDDS-001 dataset and computing AUC and ROC curves of the entire edge router. The new IDS when compared with the Random Forest classifiers attains better result in terms of accuracy and false-positive rate with mean running time of 6.23s. It achieves 97% accuracy and 0.21% false-positive rate.

In [66], the authors proposed a novel IDS method for cloud computing, based on the integration of Artificial Bee Colony, Artificial Neural Networks, and fuzzy clustering algorithm. The fuzzy clustering algorithm prepares the uniform training subsets to improve the training speed. Also, the IDS embed a Multilayer Perceptron ANN with the Artificial Bee Colony which speeds up the determination of the value of weights along with biases in the training stage of the network to distinguish between normal and abnormal traffic data. The authors used root mean square error (RMSE) with mean absolute error (MAE) as the parameters to evaluate the IDS. The proposed IDS exhibit the lowest value of RMSE and MAE with a 2.23% enhancement in correctly-classified cases over the baselines when compared with FC-ANN.

IV. CONCLUSION

In this research, we studied works that use machine learning techniques for intrusion detection, and we have provided a comprehensive survey of these methods. Studying the literature, we have identified various machine learning algorithms use in IDS and showed some of the methods adopted to optimize IDS models so as to decrease the classification time, increase the accuracy and detection rate. Particularly, we have classified IDS into two groups: single machine learning techniques and hybrid machine learning techniques in order to establish the merits and drawbacks of the members of each group. Also, we highlighted how parameter-based algorithm use in IDSs requires careful tuning of the parameters to achieve high accuracy. From our discussion, future directions can be envisioned. There is the need for further research to be carried out with the objective of achieving optimal balance between feature reduction and classification accuracy in order to enhance both the detection rate and the accuracy of detection. Also, research aimed at providing a common performance evaluation benchmark that takes into account (i) well-defined metrics and (ii) common impact factors, to make it possible to obtain a comparable assessment of different proposals is necessary.

REFERENCE

[1]. P. Mell and T. Grance (2011) The NIST definition of cloud computing.

[2]. A. Patel, M. Taghavi, K. Bakhtiyari, and J. C. JúNior (2013) An intrusion detection and prevention system in cloud computing: A systematic review. Journal of Network and Computer Applications, 36 (1): 25-41

[3]. A.L. Samuel (1967) Some studies in machine learning using the game of checkers: Recent progress. IBM J. Res. Dev. 11: 601-617.

[4]. L.E. Peterson (2009) K-nearest neighbour. Scholarpedia, 4 (2) 1883.

[5]. A. Hajimirzaei and N. J. Navimipour (2019) Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm. ICT Express, 5(1), 56-59.

[6]. M. Aljanabi, M. A. Ismail, and A. H. Ali (2021) Intrusion Detection System, Issues, Challenges, and Needs. International Journal of Computational Intelligence System, 14(1).

[7]. H. Tianfield (2017) Data mining based cyber-attack detection, System Simulation. Techno 13.

[8]. S.M.H. Bamakan, H. Wang, T. Yingjie, and Y. Shi (2016) An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization. Neurocomputing, 199:90-102.

[9]. V. Vapnik (2016) The nature of statistical learning theory. Springer science & business media, https://www.dais.unive.it/~pelillo/Didattica/Artificial%20Intelligence/Old%20Stuff/2015-2016/Slides/SLT.pdf

[10]. P. Tao, Z. Sun, and Z. Sun (2018) An improved intrusion detection algorithm based on GA and SVM. IEEE Access, 6:13624-13631.

[11]. F. Salo, M. Injadat, A.B. Nassif, A. Shami, and A. Essex (2018) Data mining techniques in intrusion detection systems: a systematic literature review. IEEE Access, 6:56046-56058.

[12]. T. Shon and J. Moon (2007) A hybrid machine learning approach to network anomaly detection. Inf. Sci.177:3799-3821.

[13]. R. Rao (2016) Jaya: a simple and new optimization algorithm for solving constrained and unconstrained optimization problems. Int. J. Ind. Eng. Comput., 7: 19-34.

[14]. G. Y. Chan, F. F. Chua, and C. S. Lee (2016) Intrusion detection and prevention of web service attacks for software as a service: Fuzzy association rules vs fuzzy associative patterns. Journal of Intelligent & Fuzzy Systems, 31(2):749-764.

[15]. K. Wang, C. Y. Huang, L. Y. Tsai, and Y. D. Lin (2014) Behaviour-based botnet detection in parallel. Security and Communication Networks, 7(11):1849-1859

[16]. M. R. Watson, A. K. Marnerides, A. Mauthe, and D. Hutchison (2016) Malware detection in cloud computing infrastructures. IEEE Transactions on Dependable and Secure Computing, 13(2):192-205.

[17]. N. C. S. Iyengar, A. Banerjee, and G. Ganapathy (2014) A fuzzy logic based defence mechanism against distributed denial of service attack in cloud computing environment. International Journal of Communication Networks and Information Security, 6(3):233-245.

[18]. X. Wang, Y. Yang, and Y. Zeng (2015) Accurate mobile malware detection and classification in the cloud. SpringerPlus, 4(1) 583.

[19]. M. T. Khorshed, A. S. Ali, and S. A. Wasimi (2012) A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing. Future Generation Computer Systems, 28 (6):833-851.

[20]. A. Dainotti, F. Gargiulo, L. I. Kuncheva, A. Pescapè, and C. Sansone (2010) Identification of traffic flows hiding behind TCP port 80. IEEE International Conference on Communications, IEEE, 2010.

[21]. E. Frank and I. H. Witten (1998) Generating accurate rule sets without global optimization. CiteSeer.

[22]. R. Lopez and E. Oñate (2006) A variational formulation for the multilayer perceptron. International Conference on Artificial Neural Networks, Springer 159-168.

[23]. G. H. John and P. Langley (1995) Estimating continuous distributions in Bayesian classifiers. Proceedings of the 11th conference on Uncertainty in artificial intelligence, 338-345.

[24]. N. Pitropakis, D. Anastasopoulou, A. Pikrakis, and C. Lambrinoudakis (2014) If you want to know about a hunter, study his prey: detection of network based attacks on KVM based cloud environments. Journal of Cloud Computing, 3(1)20.

[25]. T. F. Smith and M. S. Waterman (1981) Identification of common molecular subsequences. Journal of Molecular Biology, 147:195-197.

[26]. A. Nezarat and Y. Shams (2017) A game theoretic-based distributed detection method for VM-to-hypervisor attacks in cloud environment. The Journal of Supercomputing, 73(10)4407-4427.

[27]. O. Osanaiye, H. Cai, K. K. R. Choo, A. Dehghantanha, Z. Xu, and M. Dlodlo (2016) Ensemble-based multi-filter feature selection method for DdoS detection in cloud computing. EURASIP Journal on Wireless Communications and Networking, 2016(1)130.

[28]. J. Yu, H. Kang, D. Park, H. Bang, and D. W. Kang (2013) An in-depth analysis on traffic flooding attacks detection and system using data mining techniques. Journal of Systems Architecture, 59(10):1005-1012.

[29]. J. Peng, K. K.R. Choo, and H. Ashman (2016) Bit-level n-gram based forensic authorship analysis on social media: Identifying individuals from linguistic profiles. Journal of Network and Computer Applications, 70:171-182.

[30]. S. Rastegari, P. Hingston, and C. P. Lam (2015) Evolving statistical rulesets for network intrusion detection. Applied Soft Computing, 33: 348-359.

[31]. N. Kumar, J. P. Singh, R. S. Bali, S. Misra, and S. Ullah (2015) An intelligent clustering scheme for distributed intrusion detection in vehicular cloud computing. Cluster Computing, 18(3):1263-1283.

[32]. T. Huang, Y. Zhu, Y. Wu, S. Bressan, and G. Dobbie (2016) Anomaly detection and identification scheme for VM live migration in cloud infrastructure. Future Generation Computer Systems, 56: 736-745.

[33]. S. Sharma, A. Gupta, and S. Agrawal (2016) An Intrusion Detection System for Detecting Denial-of-Service Attack in Cloud Using Artificial Bee Colony. Proceedings of the International Congress on Information and Communication Technology, India, 2016, pp. 137-145.

[34]. M. B. and P. K. Rajendran (2015) Intelligent Intrusion Detection System for Private Cloud Environment. Proceedings of the 3rd International Symposium Security in Computing and Communications, Springer, 2015, pp. 54-65.

[35]. Z. Chiba, N. Abghour, K. Moussaid, A. E. omri, and M. Rida (2016) A Cooperative and Hybrid Network Intrusion Detection Framework in Cloud Computing Based on Snort and Optimized Back Propagation Neural Network. Procedia Computer Science, 83(2016):1200-1206.

[36]. M. Roesch (1999) Snort: Lightweight intrusion detection for networks. Lisa, 99:1.

[37]. G. Aceto, D. Ciuonzo, A. Montieri, V. Persico, and A. Pescapè (2019) MIRAGE: Mobile-app Traffic Capture and Ground-truth Creation, Proceedings of 4th IEEE International Conference on Computing Communication and Security (ICCCS 2019).

[38]. P. Ghosh, A. Saha, and S. Phadikar (2016) Penalty-Reward Based Instance Selection Method in Cloud Environment Using the Concept of Nearest Neighbour. Procedia Computer Science, 89: 82-89.

[39]. A. Chonka, Y. Xiang, W. Zhou, and A. Bonti (2011) Cloud security defence to protect cloud computing against HTTP-DoS and XML-DoS attacks. Journal of Network

and Computer Applications, 34(4):1097-1107.

[40]. K. Vieira, A. Schulter, C. Westphall, and C. Westphall (2010) Intrusion Detection for Grid and Cloud Computing. IT Professional 12:38-43.

[41]. W. Xiong, H. Hu, N. Xiong, L. T. Yang, W. C. Peng, X. Wang, et al. (2014) Anomaly secure detection methods by analysing dynamic characteristics of the network traffic in cloud communications. Information Sciences, 258:403-415.

[42]. T. Poston and I. Stewart (2014) Catastrophe theory and its applications. Courier Corporation.

[43]. H. Haken (2013) Synergetic computers and cognition: A top-down approach to neural nets. Springer Science & Business Media.

[44]. J. H. Holland (1992) Genetic algorithms, Scientific American, 267: 66-73.

[45]. B. Senthilnayaki, K. Venkatalakshmi, and A. Kannan (2015) Intrusion detection using optimal genetic feature selection and SVM based classifier, 3rd International Conference on Signal Processing, Communication and Networking (ICSCN).

[46]. J. Kennedy and R. Eberhart, Particle Swarm Optimization (PSO) (1995) Proceeding of IEEE International Conference on Neural Networks.

[47]. G. Wang, S. Chen, and J. Liu (2015) Anomaly-based intrusion detection using multiclass-SVM with parameters optimized by PSO. Int. J. Secur. Appl., 9:227-242.

[48]. M. Dorigo, V. Maniezzo, and A. Colorni (1996) Ant system: optimization by a colony of cooperating agents. IEEE Trans. Syst. Man Cybern, Part B Cybern, 26:29-41.

[49]. H.H. Gao, H.H. Yang, and X.Y. Wang (2005) Ant colony optimization based network intrusion feature selection and detection. International conference on machine learning and cybernetics, 6:3871-3875.

[50]. P. Ganeshkumar and N. Pandeeswari (2016) Adaptive neuro-fuzzy-based anomaly detection system in cloud, International Journal of Fuzzy Systems, 18:3, 367-378.

[51]. N. Pandeeswari and G. Kumar (2016) Anomaly detection system in cloud environment using fuzzy clustering based ANN. Mobile Networks and Applications, 21(3):494-505.

[52]. Y. Liu and R. Ma (2013) Network anomaly detection based on BQPSO-BN algorithm. IETE Journal of Research, 59:334-342.

[53]. S. Raja and S. Ramaiah (2016) An Efficient Fuzzy-Based Hybrid System to Cloud Intrusion Detection. International Journal of Fuzzy Systems, 19(1):1-16, February, 2016.

[54]. M.C. Su and C. H. Chou (2001) A modified version of the K-means algorithm with a distance based on cluster symmetry. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23:(6)674–680.

[55]. P. Vivekanandan, M. Rajalakshmi, and R. Nedunchezhian (2013) An intelligent genetic algorithm for mining classification rules in large datasets. Computing and Informatics, 32(1)1-22.

[56]. C. H. Tsang, S. Kwong, and H. Wang (2005) Anomaly intrusion detection using multi-objective genetic fuzzy system and agent-based evolutionary computation framework. Fifth IEEE International Conference on Data Mining (ICDM'05).

[57]. K. Shafi, and H.A. Abbass (2009) An adaptive genetic-based signature learning system for intrusion detection. Expert Systems with Applications, 36(10):12036-12043.

[58]. S. Mukkamala, A. H. Sung, and A. Abraham (2003) Intrusion detection using ensemble of soft computing paradigms. Intelligent systems design and applications, 239-248.

[59]. M. S. Hoque, M. A. Mukit, and M. N. Bikas (2012) An implementation of intrusion detection system using genetic algorithm. International Journal of Network Security and Its Applications, 4(2):109-119.

[60]. H. M. Shirazi (2010) An intelligent intrusion detection system using genetic algorithms and features selection. Majlesi Journal of Electrical Engineering, 4:1.

[61]. P. Ghosh, A. K. Mandal, and R. Kumar (2015) An Efficient Cloud Network Intrusion Detection System. Information Systems Design and Intelligent Applications, 91-99.

[62]. M. Lin and S. Chen (2015) An Efficient Anomaly Detection Framework for Cloud Computing Environment. JCP 10 (3):155-165.

[63]. X. He and P. Niyogi (2004) Locality preserving projections. Advances in neural information processing systems, 153-160.

[64]. H. Abdi and L. J. Williams (2010) Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2 (4): 433-459.

[65]. M. Idhammad, K. Afdel, and M. Belouch (2018) Distributed intrusion detection

system for cloud environments based on data mining techniques. Procedia Computer Science, 127: 35-41.

[66]. B. Hajimirzaei and N. J. Navimipour (2019) Intrusion detection for cloud computing using neural networks and artificial bee colony optimization algorithm. ICT Express, 5(1):56-59.