Machine Learning and Feature Selection Approach for Anomaly based Intrusion Detection: A Systematic Novice Approach

Amrita, Shri Kant

Abstract: ***Network Intrusion Detection System (NIDS) has become an imminent research area in network and information security due to the proliferation of the Internet and rapid increase in anomalous activities or intrusions. NIDS helps to detect anomalous activities or intrusions which compromise CIA (confidentiality, integrity, and availability), violate the security policies and mechanisms of a computer network. This paper presents a survey on anomaly based NIDS using machine learning technique employing feature selection approach. The prime contribution of this research is to present technical and empirical evaluation of each paper. The state-of-the-art NIDS is systematically analyzed and discussed according to machine learning and feature selection techniques used, number of selected features, efficiency in terms of various performance metrics and its result. This paper also provides an idea of selecting more appropriate solution and also the scope of improvement for each specific case.***

Index Terms: Anomaly Detection, Machine Learning, Network Intrusion Detection System, Feature selection.

I. INTRODUCTION

The security of computer network becomes critical due to the tremendous growth of the Internet and network attacks in recent decades. As a result, Internet based information systems become vulnerable to internal and external attack and hence Network Intrusion Detection System (NIDS) has emerged as an indispensable component. It combats the misuse, abuse, and unauthorized use of resources of computer network. In general, Intrusion Detection System (IDS) can be categorized based on detection approach for determining the occurrence of intrusions as—anomaly based or behavior based detection and misuse based or signature based detection. Anomaly based detection approach constructs the model based on normal behavior and monitor to identify deviation from the normal behavior to detect abnormal behavior [1]. This approach is able to detect novel as well as "zero days" intrusions, but often has high False Positive Rate (FPR) [2]. Whereas, misuse based detection approach identifies abnormal behavior by comparing network traffic to predefined patterns or signatures of known intrusions stored in a database [3].

Machine learning techniques have become popular in construction of NIDS as they achieve better accuracy, performance, and faster speed. An NIDS needs to analyze huge data with high dimension feature set. Redundant, irrelevant, and extra features can degrade the performance and increase the computation time of NIDSs. Hence, feature selection may be considered as preprocessing step before training of classifiers to overcome aforesaid challenges [4]. This research focuses on anomaly based NIDS, which utilizes machine learning technique employing feature selection approach evaluated on KDD [5] dataset to detect intrusion. The KDD dataset is extensively used dataset for NIDS. It consists of 4,940,000 connection instances for training dataset and 311,029 for test dataset. Each connection instance consists of 41 features plus one class label. The class label belongs to either normal or a particular attack type, which falls into one of four kinds— Probe, User to Root (U2R), Denial of Service (DoS), and Remote to Local (R2L). The details of the KDD dataset can be found in [4].

This paper presents a literature review of anomaly based NIDS published in the period of 2003-2017. Numerous anomaly-based NIDSs have been developed utilizing various machine learning techniques in the literature. This review includes single machine learning technique used in the development of NIDS employed feature selection approach evaluated on the KDD dataset. The major components analyzed during the review of each work are sole supervised machine learning technique used to develop NIDS, feature selection approach, number of selected features, feature number, performance evaluation metrics, and results reported in the literature. The main contribution of this paper is to provide technical and empirical evaluation of each surveyed paper in spite of taxonomy about NIDS, machine learning, and feature selection approach.

The remainder of the paper is as follows. The detail insight of reviewed papers on the aforementioned study is presented in the next section. Section 3 provides the scope of work and proposed methods and finally, the conclusion in Section 4.

II. A SYSTEMATIC LITERATURE REVIEW OF RELATED WORK

This section provides the state-of-art various anomaly based NIDS using the single machine learning technique, e.g. Decision Tree (DT), Genetic Algorithm (GA), k-Nearest Neighbor (k-NN), Naïve Bayes (NB), Neural Network (NN), Support Vector Machines (SVM), etc., employing a feature selection approach to construct NIDS.

Revised Manuscript Received on April 11, 2019.

Amrita, Department of Computer Science and Engineering, SET, Sharda University, Greater Noida, Uttar Pradesh, India.

Shri Kant, Department of Computer Science and Engineering, SET, Sharda University, Greater Noida, Uttar Pradesh, India.

Each surveyed paper is analyzed and discussed based on what machine learning and feature selection technique is used to construct NIDS, number of feature and features selected by employed feature selection technique, performance evaluation metrics, e.g. classification Accuracy (ACC), Detection Rate (DR), True Positive Rate (TPR), False Negative Rate (FNR), False Alarm Rate (FAR), True Negative Rate (TNR), FPR, etc., used to measure the performance of NIDS and their results.

A class-specific detection based on SVM for IDS is presented in [6]. Two methods for feature ranking are applied to identify the number of important features for each of 5 classes (Normal, DoS, Probe, U2R, and R2L) as shown in Table 1. The system has most remarkable performance: the training and testing time decrease for each class; the ACC decreases little for classes DoS, Probe, and R2L, and remains equal for 'Normal' and U2R as shown in Table I.

The stochastic radial basis function NN output sensitivity measure is proposed in [7], which is based on the feature importance ranking method using sensitivity measure. It is employed to assess the features for Normal and DoS attack only. The 8 most sensitive features {2, 23, 24, 29, 32, 33, 34, and 36} are sufficient to detect Normal and DoS. It is found that ACC can be maintained at the same level with only 8 features with less computation complexity, which is decreased from 23 to 9 seconds. The ACC of DoS and Normal are 99.06% and 99.77% respectively. The FAR for 8 features (41 features) are 0.18% (0.01%) and 0.27% (0.03%) and the FPR are 0.93% (0.70%) and 0.94% (0.71%) in training and testing respectively.

A simple GA using k-NN classifier as a fitness function is utilized in [8]. Features are then ranked based on weighted feature set evolved by GA for attack classes. The selected top 5 ranked features are as DoS{1, 11, 23, 24, and 29}, Probe{2, 3, 6, 30, and 37}, R2L{3, 12, 23, 24, and 36}, and U2R{6, 17, 24, 31, and 41}. The result demonstrates that there is an increase in ACC by employing weighted feature sets.

The classifiers Back Propagation NN (BPNN) and SVM are used to evaluate the proposed feature selection algorithm, Feature Selection Method based on Davies–Bouldin Index, in [9] for 5-class (Normal, DoS, Probe, R2L and U2R). In this, single best feature set consists of 24 features {1, 3, 4, 5, 6, 8, 10, 11, 12, 13, 23, 24, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 38, and 39} are selected for 5-class. The ACC of classifiers BP network and SVM employing this feature set are 0.1017 and 0.056 respectively.

BPNNs are utilized as a classifier and ReliefF Immune Clonal GA as feature subset selection is proposed in [10]. ReliefF Immune Clonal GA is a combination of ReliefF algorithm, Immune Clonal selection algorithm and GA. Experimental results exhibit that proposed method has better ACC than GA and ReliefF-GA on selected features of size 8. It has higher ACC (86.47%) than ReliefF-GA.

The Least Square SVM (LSSVM) classifier and Ant Colony Optimization (ACO) for feature selection approach are presented in [11]. The number of features selected is 9 for DoS, Probe, and U2R & R2L are 9, 11, and 14 respectively. The experiments exhibit that feature selection approach not only decrease the number of features but also boost the performance of classifier and make detection more effective in terms of time. The experimental results in terms of ACC, FPR and average detection time (ADT) (ms/sample) for attack types are as DoS (95.2, 3.24, 0.031), Probe(99.4, 0.35, 0.074) and U2R & R2L(98.7,1.60, 0.078) respectively.

A combination of GA, SVM, and Correlation-based Feature Selection (CFS) is proposed in [12]. The DR and FPR in average are 99.56% and 37.5% using 12 selected optimal features {1, 6, 12, 14, 23, 24, 25, 31, 32, 37, 40, and 41}. The experiment shows that the DR of selected feature set is lower than full features set (the difference is around 0.83% in average). But there is decrease in training and testing time significantly though retaining the DR and FPR within acceptable range.

SVM and Discriminant Analysis are combined for anomaly based NIDS to identify network intrusion in [13]. Nine features {2, 12, 23, 24, 29, 31, 32, 36, and 39} are obtained by Discriminant Analysis and assessed by SVM. The TPR, TNR, FPR, and FNR of proposed method are 90.07%, 99.58%, 0.42%, and 9.93% respectively.

Resilient BPNN as a classifier and normal distribution, beta distribution, chi-square analysis, and logistic regression based on ranking and feature selection approach is utilized in [14]. Features are ranked based on their influence towards the final classification by using NN employing forward selection and backward elimination methods. The selected 25 features {1, 2, 3, 5, 8, 10, 12, 13, 22, 24, 25, 26, 27, 28, 29, 30, 34, 33, 35, 36, 37, 38, 39, 40, and 41} are ranked by chi-square test to classify the network traffic according to 5-class (Normal, Probe, DoS, U2R, and R2L). Experiments show that resilient BPNN exhibits high ACC and needs less training and testing time than classical NN. The overall ACC is 97.04 with FNR of 0.20 and FPR of 2.76.

The effectiveness of Rough Set theory as a classifier and also to identify the important features to build IDS is examined in [15]. The classification ACC using obtained 6 features {3, 4, 5, 24, 32, and 41} for Normal, DoS, Probe, R2L, and U2R are 89.84, 99.34, 99.63, 100, and 100 respectively. The results demonstrate that obtained feature subset is robust and has reliable performance.

A lightweight IDS based on Classic Maximum Entropy model is proposed in [16]. Chi-Square and Information Gain (IG) are used to select significant features. Experimental results show that the proposed model is effective and have

Table I: Performance of all and selected features

Class	# & Features	Training Time (sec)		Testing Time (sec)		ACC (%)	
		All	Selected	All	Selected	All	Selected
Normal	20 {1-6,10,12,17,23,24,27, 28,29,31-34,36,39}	7.66	4.58	1.26	0.78	99.55	99.55
Probe	11 {1-6,23,24,29,32,33}	49.13	40.56	2.10	1.20	99.70	99.36
DOS	11 {1,5,6,23-26,32,36,38,39}	22.87	18.93	1.92	1.00	99.25	99.16
R2L	6 {1,3,5,6,32,33}	11.54	6.79	1.02	0.72	99.78	99.72
U2R	10 {1-6,12,23,24,32,33}	3.38	1.46	1.05	0.70	99.87	99.87

good ACC especially for DoS attack with reduced testing time. The ACC and testing time on selected 12 features {3, 5, 6, 10, 13, 23, 24, 27, 28, 37, 40, and 41} are as Normal (99.73, 0.78s), DoS (100, 1.03s), Probe (99.76, 1.25s), U2R (99.87, 0.70s), and R2L (99.75, 0.68) respectively.

Lee et al. [17] utilized Minimax Probability Machine employing Random Forests (RF) as feature selection for DoS attack only because other attack types have very less number of records and not suitable for experiments. RF ranked features by numeric values, so top 5 important features {3, 5, 6, 23, and 29} are selected, which exhibits DR as 99.84% and 0.1039 sec for average simulation time. The experimental results show that this approach is superior to the previous approaches.

SVM classifier with Radial Basis Function network kernel employing Decision Dependent Correlation as feature selection method is proposed in [18]. Top 20 features {2, 3, 5, 7, 8, 9, 10, 11, 13, 14, 15, 17, 18, 22, 24, 27, 28, 36, 40, and 41} are selected by calculating mutual information and decision of each feature. The ACC of the proposed system is 93.46%. This method is compared with Principal Component Analysis (PCA) and it outperforms PCA.

A misuse detection system is proposed by [19] by investigating the possibility to increase DR of U2R attack. PCA and Multi Expression Programming are employed to extract features and GA is utilized to implement rules to detect types of attacks. The number and selected features for various attacks are as U2R 2{14 and 33}, DoS 3{1, 5, and 39}, and Normal 3{3, 10, and 12} respectively. The experimental results show that this system performed better than the best-performed model reported in the literature.

SVMs employing modified random mutation hill climbing feature selection method based on wrapper approach for attack classes is presented in [20]. The experiments exhibit that system has higher DR of detecting known as well as new attacks on obtained features than all features with decrease in training and testing time as shown in Table II.

Table II: Performance on selected and all features

Class	#Features	Selected Features	Avg. Train Time		Avg. Test Time	
			All	Selected	All	Selected
All	5	3,5,23,33,34	78	30	18	6
DOS	4	5,12,23,34	136	31	22	5
Probe	5	1,3,5,23,37	245	96	49	17
R2L	3	1, 5,6	317	24	55	7
U2R	5	1,3,6,14,33	193	78	50	15

A Multi-Objective Genetic Fuzzy IDS is proposed in [21]. It is used to search near-optimal feature subset. The selected

feature subset decreases the computational effort and enhances the performance of the system. The selected 27 features {2, 5, 6, 7, 8, 9, 11, 12, 13, 14, 17, 18, 22, 23, 25, 30, 32, 33, 34, 35, 36, 37, 38, 39, and 40} demonstrate that the proposed approach produces lowest FAR (1.1%) and highest ACC (99.24%) with minimum number of features in the paper.

C4.5-PCA-C4.5, a new hybrid approach, is proposed in [22]. It uses C4.5 as a classifier and PCA and DT (C4.5) as feature selection methods. This approach has the lowest FPR, highest TPR, fast training and testing process using only 7 important extracted features {1, 3, 4, 10, 22, 33, and 34}.

Panda and Patra [23] proposed a framework of NIDS based on NB. The dataset is grouped into four attack types (DoS, Probe, R2L, and U2R). The classifier NB has achieved ACC of 99%, 96%, 90%, and 90% respectively on attack types, average FAR of 3%, DR of 95%, and with error rate 5%. The proposed framework performed faster (1.89 seconds) to build the model, efficient and cost effective.

Sheen and Rajesh [24] proposed DT employing different feature selection methods—IG, Chi square, and ReliefF. Top 20 selected features are {2, 3, 4, 5, 12, 22, 23, 24, 27, 28, 30, 31, 32, 33, 34, 35, 37, 38, 40, and 41}. The experiments show that both IG and Chi square have similar performance while ReliefF has lower performance. The ACC of IG, Chi Square, and ReliefF are 95.85%, 95.85%, and 95.64% respectively.

Wang et al. [25] utilized Bayesian networks (BN) and DT (C4.5) for classifier, and filter and wrapper approaches for feature selection. BN with filter based IG approach and DT with wrapper based approach are employed to select subsets of features. Ten features are selected for each attack class by using IG, BN, and DT. The empirical results show that DR and FPR of classifiers BN and DT using only 10 features remains almost the same or even becomes better compared to all 41 features with reduced training and testing time as shown in Table III.

Three machine learning techniques—DT, Particle Swarm Optimization (PSO) and Flexible Neural Tree (FNT) utilized by [26] for feature selection. The 5 important features {10, 11, 13, 14, and 17} are obtained based on the involvement of the features to build the DT. The results demonstrate that DT gives better DR, FPR, and cost per example than FNT and PSO for Normal, DoS, Probe and cost per example than FNT and PSO for Normal, DoS, Probe

Table III: Performance using 41 features and selected 10 features

Attacks	Selected Features	Methods	Using 41 Features				Using 10 Features			
			DR	FPR	Train Time(s)	Test Time(s)	DR	FPR	Train Time(s)	Test Time(s)
DoS	3,4,5,6,8,10,13, 23,24,37	BN	98.73	0.08	04.7	2.1	99.88	0.00	0.8	0.6
		C4.5	99.96	0.15	16.3	1.2	99.87	0.14	4.6	0.5
Probe	3,4,5,6,29,30,32,35,39 ,40	BN	92.89	6.08	03.1	2.8	82.93	3.06	0.5	0.4
		C4.5	82.59	0.04	14.5	1.1	82.88	0.05	1.2	0.3
R2L	1,3,5,6,12,22,23,31,32 ,33	BN	92.22	0.33	02.6	1.8	89.33	0.32	0.5	0.4
		C4.5	80.29	0.02	10.5	0.8	87.34	0.01	0.5	0.2
U2R	1,2,3,5,10,13,14, 32,33,36	BN	75.86	0.29	02.6	1.8	65.5	0.12	0.4	0.4
		C4.5	24.14	0.00	09.9	0.7	24.14	0.00	0.6	0.2

and R2L, but U2R. DRs of DT are as Normal (99.96), DoS (100), Probe (99.66), R2L (99.02), and U2R (88.33).

The work in [27] proposed to sample different ratios of normal data to achieve better ACC and to compare the efficiency of machine learning methods DT (C4.5) and SVM in IDS. The performance of DT and SVM are compared with KDD winner and found that DT is better than SVM in ACC and DR, but SVM is superior in FAR. The DRs of DT are as DoS (62.96), Probe (86.30), U2R (50.06), and R2L (17.43) and average FAR is 1.44.

An approach utilizing Artificial NN (ANN) and SVM to detect an attack is proposed in [28]. The DR of ANN are as DoS (59.1), Probe (82.4), U2R (65.9), and R2L (14.3) and for SVM as DoS (63.1), Probe (83.8), U2R (66.3), and R2L (14.9). The result exhibits that SVM is superior to ANN.

A Quantitative Intrusion Intensity Assessment (QIIA) approach is presented in [29], which utilizes two methods QIIA1 and QIIA2 to find the value of threshold parameters. The top 5 significant features {3, 6, 10, 23, and 32} are selected by QIIA by utilizing RF to detect only DoS attacks. The QIIA1 and QIIA2 achieve DR of 97.94 and 99.37 respectively. Other attack types are not considered as they have very less number of records.

An IDS utilizing fuzzy association rule mining to build stateless classifiers to classify normal and attack is proposed in [30]. The performance of their IDSs build for misuse detection and anomaly detection are compared. The IDS obtains DR (80.6%) and FPR (2.95%) for anomaly detection, compared with DR (91%) and FPR (3.34%) for misuse detection.

Xiao et al. [31] proposed two-step feature selection approach assessed by C4.5 and SVM. The obtained 21 features {1, 3, 4, 5, 6, 8, 11, 12, 13, 23, 25, 26, 27, 28, 29, 30, 32, 33, 34, 36, and 39} demonstrate that DR and FAR increase little on selected features and the processing speed improves to 30.73%. The DR, processing time, and FAR of obtained features (all features) are 86.30(87.00), 15.16(21.89) sec and 1.89(1.85) respectively.

Robust Artificial Intelligence Selection, a hybrid approach, is proposed in [32]. It is based on SVM and mutual information feature subsets selection. The training and testing times for the classifier are decreased by using reduced features and has the lowest FAR (3.49%), the highest ACC (99.01%), and DR (99.27%).

Lee and He [33] presented two-stage entropy-based traffic profiling method to identify network attack (only DoS). Only 23 features (basic and time-based) from KDD dataset are considered for feature selection. The top 6 selected features {5, 6, 31, 32, 36, and 37} are ranked based on ACC. The experimental results show that this method achieved lower complexity and superior TPR of 91%.

A novel approach, Enhanced Support Vector Decision Function (ESVDF), is proposed by [34]. This approach utilizes support vector decision function to rank the features and Backward Elimination Ranking (BER) or Forward Selection Ranking (FSR) methods to correlate the features. Nine and 6 features are selected by ESVDF/BER and ESVDF/FSR respectively and evaluated using two classifiers—NN and SVM. The empirical results exhibit that SVM performed better than NN in ACC and training time. The ACC of ESVDF/BER and ESVDF/FSR by SVM on selected features are 99.58 and 99.46 and FPR are 0.0031 and 0.0033 respectively.

A novel approach, which employed BN and two feature selection approaches as consistency subset evaluator and CFS subset evaluator, is proposed in [35]. The ACC of proposed system for Normal, Probe, R2L, DoS, and U2R attack types using 7 selected features {3, 6, 12, 23, 32, 14, and 40} are 99.8, 89.4, 91.5, 99.9, and 69.2% respectively.

Two schemes—known detection scheme and unknown detection scheme are proposed in [36]. The known detection scheme employed DT (C5.0) with Euclidean Distance whereas the unknown detection scheme employed DT (C5.0) with Cosine Similarity are utilized to select features for known and unknown attacks respectively to construct a model. The known and unknown detection schemes extract 30 {1, 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 22, 25, 26, 27, 28, 30, 31, 35, 37, 38, 39, 40, and 41} and 24 {1, 2, 3, 4, 12, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 35, 36, 37, 38, 39, 40, and 41} important features respectively. The TPR, FPR, time to build model (sec) using selected features are for known attack (98.12, 1.87, 51s) and for unknown attack (68.28, 31.72, 45s) respectively.

A lightweight NIDS using new hybrid feature selection approach is proposed in [37]. This approach uses enhanced C4.5 and Chi-Square approaches for feature selection. The top 5 extracted features by C4.5-Chi2 are {3, 4, 5, 8, 25}. The experimental results demonstrate that there is significant decrease in training (0.02 sec) and testing time (0.03 sec) while retaining high DR and low FPR as Normal

(99.9,1.6), Probe (93.87, 1.82), R2L (61.55, 12.17), DoS (99.3, 1.48), and U2R (50.01, 28.32) respectively.

An anomaly detection based random effects logistic regression model is proposed in [38], which not only considers system characteristics, but also the uncertainty that cannot be explained by such predictor characteristics. As a result, five input variables are selected as (2, 10, 12, 13, and 24). The ACC of the proposed model on training dataset is 98.96%, while on validation data set is 98.74%.

An anomaly detection based approach is proposed in [39], which utilizes distributed time-delay ANN. The training dataset contains 25000 instances (5000 instances for each class of Normal, DoS, U2R, Probe, R2L), and testing dataset contains 2500 instances (500 instances for each type) are used. The experimental results show that overall ACC is 99.88%. The ACC for all classes are as: Normal (98.40%), DoS (97.60%), U2R (96.20%), Probe (98.20%), and R2L (95.80%).

An automatic feature selection method using CFS based on filter approach, evaluated by BayesNet and C4.5, is proposed by [40]. The number of features and selected features are as—for Normal&DoS are 3{5, 6, and 12}, Normal&Probe are 6{5, 6, 12, 29, 37, and 41}, Normal&R2L are 2{10 and 22}, and Normal&U2R is 1{14}. Average ACCs of BayesNet and C4.5 are 98.82% and 99.41% respectively. This method outperforms the GA-CFS and best-first-CFS methods.

A novel inconsistency-based feature selection method with DT (C4.5) is proposed by [41]. The proposed method is compared with CFS and it outperforms CFS as shown in Table IV. This method is simple and quick and can be applied for lightweight IDS.

A features extraction based customized features to enhance the ACC of the signature detection classification model is proposed in [42]. Eleven features are selected and experiment is carried out employing three randomly selected datasets from KDD and four data mining methods— DT, PART, Ridor, and Ripper (Jrip). The result shows that DRs have been increased between 0.4% to 9% and FARs decreased between 0.17% to 0.5%.

BP based ANN is proposed in [43] for IDS for the classification of normal and attack. 2570 records are selected from KDD dataset, of which 1325 for training (Normal=631, Attack=694) and 1245 for testing (Normal=523, Attack=722). The experimental results are as DR (80.5%), FAR (7.4%) and FNR (11.3%).

A fuzzy class-association rule mining approach is proposed in [44] based on genetic network programming utilizing sub-attribute to detect intrusion for NIDS. This approach can be used for anomaly as well as misuse detection. The number of features and selected features are not mentioned in the paper. The DR, FPR, and FNR are 98.7, 0.53, and 3.75 respectively.

A method for IDS utilizing NB and an improved IG method based on feature redundancy is presented in [45]. Twelve features {2, 3, 5, 6, 8, 10, 12, 23, 25, 36, 37, and 38} are selected by applying improved IG. The experiments are performed on 41 features and 12 features exhibit DR of 96.19 and 96.80; processing times (in sec.) of 8.34 and 2.08; and FPR of 5.22 and 1.02 respectively.

A novel approach, genetic quantum PSO, is proposed in [46] to reduce features for NIDS. The empirical result demonstrates that this method is more efficient than quantum PSO and PSO methods to eliminate redundant and independent features. DR and speed of NIDS are greatly increased by employing this method evaluated by SVM as shown in Table V.

A method employed GA combined with weighted k-NN as the fitness function for feature selection to detect only DoS attacks for anomaly based NIDS in real time is proposed in [47]. The best number of features and optimal weight vector of features with their weights are selected. The top 19 features for known attacks and the top 28 features for unknown attacks are considered for NIDS. An overall ACC of 97.42% and 78% are achieved for known and unknown attacks respectively. Selected features are not listed in the paper.

Hidden NB approach, which relaxes the NB's assumption of conditional independence, is proposed in [48]. The feature set {3, 5, 6, 12, 23, 31, and 32} consists of 7 out of 41 features. The empirical results demonstrate that hidden NB approach achieved overall a superior performance in terms of ACC, misclassification cost, and error rate compared to classical NB approach and other leading state-of-the-art models. This model has ACC (93.73) and Error rate (6.28).

A reliable and efficient classifier is built by [49] to classify network traffic to be normal or not. It is a combination of SVM, ACO and clustering method. Nineteen critical features {2, 4, 8, 10, 14, 15, 19, 25, 27, 29, 31, 32, 33, 34, 35, 36, 37, 38, and 40} are selected by employing gradually feature removal approach. SVM achieved ACC of 98.62%, average Matthews correlation coefficient of 0.8612 with greatly reduced training and testing time.

The paper [50] proposed Bees algorithm (BA) using SVM, a wrapper-based feature selection method. BA is used as a search strategy in wrapper method for subset generation. The experiments utilized four random subsets collected from KDD. The result shows that the feature subset of 6 features {3, 12, 24, 25, 32, and 37} produced by BA-SVM has yielded ACC of 93.30 and DR of 95.75.

A new hybrid IDS is proposed in [51]. It uses simplified swarm optimization for classification of intrusion data and intelligent dynamic swarm based rough set for feature selection. The selected 6 most relevant features {3, 5, 6, 27, 33, 35} achieved higher ACC of 93.3% than the classical PSO, SVM, and NB classifiers.

A lightweight IDS utilizing neurotree is proposed in [52] for multi-class classification to detect anomalies in networks. A wrapper based feature selection approach is employed, which reduces classifiers' computational complexity with great impact. The features extracted by this

Table IV: Performance of all, proposed and CFS method

Attack Type	All features		Proposed Method			CFS Method		
	ACC(%)	Train Time(s)	# & Features	ACC(%)	Train Time(s)	# & Features	ACC(%)	Train Time(s)
All	99.50	3.72	8(1,3,5,25,32, 34,36,40)	99.45	0.48	11(2,3,4,5,6,10,23,24,25,36,37)	99.67	6.28
DoS	99.94	1.08	4(3,4,10,23)	99.81	0.22	4(2,5,16,22)	99.32	0.33
Probe	99.85	0.66	4(3,5,35,36)	99.77	0.16	4(5,6,25,37)	94.35	0.27
R2U	98.99	0.22	5(3,5,12,32,35)	99.13	9.13	5(3,5,10,24,33)	98.05	0.11
U2R	100.00	0.11	2(3,41)	100.00	0.09	9(3,10,24,29,31,32,33,34,40)	100.00	0.08

Table V: Performance of SVM using selected and all features

Attack Type	# & Features	Training Time(ms)		Detecting Time(ms)		DR		Error Report Rate(%)	
		Selected	All	Selected	All	Selected	All	Selected	All
DoS	10 (2,6,3,12,21,22,31,26,28,30)	0.0627	0.261	0.0581	0.486	99.98	96.40	0.00	0.0013
Probe	5 (5,12,26,32,34)	0.0431	0.270	0.0478	0.164	91.77	58.90	0.001	0.00
R2L	7 (10,23,25,29,26,33,35)	0.0530	0.274	0.0140	0.352	98.26	81.47	0.00	0.0016
U2R	5 (2,3,17,32,36)	0.0006	0.001	0.0016	0.035	100	66.70	0.0003	0.0012

method are 16 {2, 3, 4, 5, 6, 8, 10, 12, 24, 25, 29, 35, 36, 37, 38, and 40}. The proposed method achieved DR of 98.4% which is superior to other methods like C4.5, Decision Stump, NB, RF, Representative tree, and Random Tree.

A novel network anomaly detection approach is presented in [53]. It identifies attack features to detect previously unknown attacks. The DT (C4.5) classification, NB feature selection, and k-means clustering methods are uniquely combined by effects-based feature detection method. The ACC of the proposed approach is 99.95.

A hybrid filtering feature selection approach is proposed in [54]. This approach removes useless and irrelevant features by selecting and ranking reliable features for more reliable and accurate IDS. Two filter based approaches— symmetrical uncertainty and IG are employed to create two reliable feature subsets in the first phase. Then these subsets are fused, weighted and ranked to get the significant features. The selected 4 important features for each group are as Normal-Dos= {2, 5, 23, and 36}, Normal-Probe= {4, 5, 27, and 29}, Normal-U2R={10, 13, 14, and 17} and Normal-R2L = {3, 5, 10, and 33}. The DR and FPR are as Normal-Dos (98.8, 0.01), Normal-Probe (95.8, 0.027), Normal-U2R (99.85, 0.002), and Normal-R2L (98.72, 0.005) respectively.

A new hybrid feature selection method is proposed by [55] and its performance is measured by the classifiers NB and C4.5. Six features {3, 5, 6, 10, 13, and 29} are obtained using proposed method and classifiers NB and C4.5 yield ACC of 99.4% and 99.9% respectively. This method outperformed standard feature selection methods and 41 features dataset on various performance metrics.

In paper [56], new approach Combining Support Vectors with Ant Colony (CSVAC), which combines SVM with Clustering based on Self-Organized Ant Colony Network (CSOACN) to get the benefits of both while ignoring their limitations, is proposed. This approach achieved DR (94.86), FPR (6.01) and FNR (1.00). Experiments demonstrate that CSVAC outperforms CSOACN alone and SVM alone in terms of DR and faster run-time.

The research proposed in [57] utilized SVM as a classifier, and PCA and GA as a feature selection approach. Genetic principal components (GPC) are obtained by searching the PCA space by applying GA to obtain features subset. The proposed method achieved DR (99.96) using 10 GPC and DR (99.94) using 12 GPC.

A new robust algorithm, BA-Membrane Computing, is proposed in [58], to improve the BA for feature subset selection using SVM. The selected 10 features (2, 3, 8, 13, 20, 24, 32, 37, 39, and 40) produced very high DR of 89.11%, ACC of 95.60% and FAR of 0.004 compared to other approaches listed in the paper.

A novel multi-layer SVM model combining kernel PCA with GA for intrusion detection is presented by [59]. Kernel PCA is used to get the principal features from data. Radial basis kernel function based on Gaussian kernel function is built to cut down the training time and enhances the performance of SVM. GA is employed to choose appropriate parameters for SVM. This model exhibits DR of 94.22 and FAR of 1.03. By comparison with other detection algorithms, the empirical results show that this model has higher ACC, better generalization and faster convergence speed compare to other detection algorithm.

The paper [60] proposed a new feature-selection approach based on the cuttlefish optimization algorithm and DT classifier as a judgment on the selected features. The proposed model obtained ACC (91.99), DR (91.00), FPR (3.92) using 5 features and ACC (73.27), DR (71.09), FPR (17.69) with 41 features. The results demonstrate that reduced feature subset provides higher DR and ACC with a lower FAR compared to all 41 features.

A hybrid method employing SVM and GA is proposed for intrusion detection in [61]. It is used to reduce the number of features. With obtained 10 features {2, 3, 4, 8, 17, 22, 23, 31, 34, and 36}, the proposed method is capable of attaining TPR of 97.3 and FPR of 1.7.

A method, ACO-FS-SVM, is proposed in [62] for network intrusion detection. It combines ACO with SVM, in which ACO is used to identify the features by means of feature weighting SVM. The experimental results

demonstrate that this method can efficiently reduce the number of features as Normal=13{2, 3, 4, 7, 8, 9, 10, 15, 16, 21, 22, 23, and 25}, DoS=10{2, 3, 7, 9, 16, 20, 27, 32, 37, and 40}, Probe=9{2, 4, 9, 21, 29, 32, 33, 34, and 35}, U2R=11{2, 4, 9, 20, 31, 21, 29, 32, 33, 34, and 35}, and R2L=13{1, 2, 3, 4, 6, 7, 9, 11, 16, 20, 21, 23, and 27} with DR as Normal (99.13), DoS (97.09), Probe (98.46), R2L (98.56), and U2R (98.68).

A new method based on feature average of total and each class is proposed in [63] for feature selection to build IDS. The selected features obtained by the proposed method is evaluated by DT are computationally effective and efficient. The result demonstrates that with 22 features, the system achieved the highest ACC (99.79%) in comparison with the ACC (99.76) on full data and other standard feature selection methods on NSL-KDD dataset [64].

A hybrid approach, Mutual Information-Binary Gravitational Search Algorithm (MI-BGSA), is presented in [65] for feature selection. MI based on filter approach is integrated into wrapper based BGSA to obtain the features, which is evaluated by SVM and tested on NSL-KDD dataset. The proposed approach selects 5 features {3, 4, 5, 6, 25} and achieved higher ACC (88.36) and DR (86.31) and low FPR (8.89) as compared to some standard filter based and wrapper based feature selection approaches.

A GA based technique is proposed to detect intrusion for network in [66]. GA is used to discover a set of simple, interval-based rulesets. This system achieved ACCs of 76.2 using 8 features and 9 rules; 75.9 using 15 features and 18 rules; 78.0 using 32 continuous features and 19 rules on NSL-KDD dataset.

A new IDS is proposed in [67] by using Layered Approach (LA) based classification algorithm utilizing Conditional Random Field (CRF) for feature selection. The optimized number of features selected are as DoS=5 {23, 34, 38, 39, and 40}, Probe= 5 {1-5}, R2L=11 {1, 5, 10, 11, 12, 13, 17, 18, 19, 21, and 22}, and U2R=11{1, 5, 10, 11, 12, 13, 17, 18, 19, 21, and 22} using CRF. The proposed system provides high ACC as Probe (98.83), DoS (97.62), U2R (86.91), and R2L (32.43).

An IDS, LSSVM-IDS-Flexible Mutual Information Feature Selection (LSSVM-IDS-FMIFS), is proposed in [68]. The FMIFS selected 17 features as {1, 2, 3, 4, 8, 10, 11, 12, 19, 23, 24, 25, 29, 31, 32, 36, 39}. The proposed system achieved ACC (99.79), DR (99.46), and FPR (0.13).

A new IDS is presented in [69], which combines an optimal Feature Selection (FS) algorithm based on IG Ratio and two classifiers SVM and Rule Based Classification. The optimal FS algorithm has selected 10 important features and achieved ACC as DoS (99.25), U2R (96.00), Probe (96.16), and R2L (96.00) by SVM.

ACO for feature selection and SVM as classifier for IDS have been proposed by [70]. The Results reveal that TPR achieved is 98.00 is significantly improved with 14 features.

The paper [71] proposed NIDS, in which it uses three methods. First, the entropy-based feature selection is utilized to obtain significant features. Second, fuzzy rules are generated by employing fuzzy control language and finally, layered classifier is developed to identify different network attacks. The system is evaluated on 10%, Whole and Corrected dataset of KDD. The overall results for DR, Recall, and FPR on 10% dataset are 98.49, 98.50%, and 1.41; on Whole dataset are 98.65, 98.47%, and 1.35; and on Corrected dataset are 99.16, 99.03%, and 0.74 respectively. The DR and FPR are significantly improved for various attacks compared with various other approaches.

An enhanced model to increase attacks detection ACC and to improve overall system performance is proposed in [72] for feature selection. This model selects 12 most pertinent features {1, 3, 5, 6, 10, 14, 23, 27, 33, 35, 36, and 38} and able to correctly detect traffic instances of Normal (99.97%), DoS (99.98%), Probe (99.3%), R2L (98.1%), and U2R (72.22%). The results indicate that 12 features have almost the same performance as of the 41 full features.

A new and effective framework, Logarithm Marginal Density Ratios Transformed Data with SVM (LMDRT-SVM), is proposed in [73]. LMDRT is used to transform original features into better quality features to enhance the detection capability of SVM. The experimental results demonstrate that it attains more robust and better performance in terms of ACC (99.93), DR (99.94), FAR (0.10), and training speed compared to existing other new methods.

III. SCOPE OF THE WORK AND PROPOSED METHOD

This survey provides many insights in the field of NIDS, machine learning techniques and feature selection approaches to detect intrusions (attacks) and normal traffic. This section discusses about the scope of the work which needs further research and proposes number of methods based on these insights.

The key challenges for NIDS to be robust, accurate, and efficient. There is always a scope to enhance the ACC of detection with high degree of confidence to differentiate normal and intrusive network traffic. From literature survey, it can be observed that none of single machine learning technique is complete and can detect intrusion from network traffic with high ACC and minimum FAR. So, using single classifier to construct robust and efficient NIDS may not be a good solution. Therefore designing and developing more robust classifiers by using multiple or ensemble or hybrid classifiers are needed. This approach may further enhance the performance of NIDS.

Network data always contain huge and high-dimensional feature set, which include redundant or irrelevant features. This makes NIDS near to impossible to work in real time and can detect the intrusion accurately with less computational effort (less detection time). The optimal number of effective features can reduce the processing time (training and testing time) and contribute to construct lightweight NIDS suitable for real time with high ACC and low FAR. There are many

feature selection approaches reported in the literature having their own advantages and disadvantages. So, selecting specific or appropriate one, which can perform well for given data and classifier is very difficult. It is recommended to combine two or more feature selection approaches and utilized them as single hybrid approach may improve its performance in general and make NIDS to work in real time.

Another possible area of excel in NIDS design is to detect all type of attacks (DoS, Probe, R2L, and U2R), known as multi-class-classification, with high ACC. From the literature survey, it can also be observed that none of single strong machine learning technique is complete and can able to do this. This type of problem needs different solution rather than classic classification paradigm. Therefore, it is essential to foresee an efficient method like multi-class/ensemble/hybrid classifiers or binarization of multi-class or specific feature set for each attack type or combination of these methods to detect each attack type with high ACC and low alarm rates.

Abundant literature is presented and numerous approaches have been proposed on NIDS. In spite of significant progress, there are still a lot of opportunities to enhance the performance of NIDS. Based on insights in the aforesaid field and scope of the work, we propose the following methods:

First, hybrid approach for feature selection based on two or more combination of different types of feature selection approaches, which is able to obtain optimal number of effective features and able to maximize the ACC, minimize the FPR, FNR, training and testing time of the system in which it is being used. This method will be useful in any domain in which there is a need for feature selection method.

Second, a robust classifier based on ensemble of diverse machine learning techniques with the aim to improve the ACC of the NIDS by utilizing the strength of each classifier and overcome the limitation of each one.

Third, develop lightweight, robust, accurate, and efficient NIDS by utilizing the second proposed method and incorporating obtained features from the first proposed method.

IV. CONCLUSION

Network Intrusion Detection System (NIDS) plays a very vital role in network security to protect the network against intrusion or attacks. Machine learning techniques have established as an active research topic to construct efficient NIDS. The redundant and irrelevant features of dataset often degrade the performance of machine learning technique in turn NIDS. Therefore employing feature selection techniques become important to reduce features and enhance the performance of NIDS.

In this paper, the survey on anomaly based Network Intrusion Detection System (NIDS) is presented. In particular, the research papers based on machine learning technique employing feature selection approach to construct NIDS are considered for this study from technical and practical aspects. The main contribution of this paper is to systematically analyze and discuss each paper, which helps to identify existing research challenges and future research directions.

Finally, the research challenges in terms of scope of work and proposed methods have been forecasted to provide further research directions to build lightweight, robust, accurate, and efficient NIDS. We have also been working on the proposed methods and finding good results.

REFERENCES

1. M. H. Bhuyan, D. K. Bhattacharyya, and J. K. Kalita, “Network anomaly detection: methods, systems and tools”, *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 303-336, 2014.
2. D. Joo, T. Hong, and I. Han, “The neural network models for IDS based on the asymmetric costs of false negative errors and false positive errors”, *Expert Systems with Applications*, vol. 25, no. 1, pp. 69-75, 2003.
3. H. Wu and S. S. Haung, “Neural network based detection on stepping stone intrusion”, *Expert System with Applications*, vol. 37, pp. 1431-1437, 2010.
4. Amrita and P. Ahmed, “A study of feature selection methods in intrusion detection system: a survey”, *International Journal of Computer Science and Engineering and Information Technology Research (IJCSEITR)*, vol. 2, no. 3, pp. 1–25, 2012.
5. KDD Cup 1999 Dataset. (1999) Available on: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html.
6. S. Mukkamala and A. H. Sung, “Feature selection for intrusion detection using neural networks and support vector machines”, *Journal of the Transportation Research Board of the National Academics*, Transportation Research Record No 1822, pp. 33-39, 2003.
7. W. W. Y. Ng, R. K. C. Chang, and D. S. Yeung, “Dimensionality reduction for denial of service detection problems using RBFNN output sensitivity", In Proceedings of the International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693), 2003, pp. 1293-1298.
8. M. J. Middlemiss and G. Dick, “Weighted feature extraction using a genetic algorithm for intrusion detection”, The Congress on Evolutionary Computation, CEC’03, 2003, pp. 1669-1675.
9. L. Zhang, G. Sun, and J. Guo, “Feature selection for pattern classification problems”, The 4th International Conference on Computer and Information Technology (CIT’04), 2004.
10. Y. Zhu, X. Shan, and J. GuO, “Modified Genetic Algorithm based Feature Subset Selection in Intrusion Detection System”, In Proceedings of IEEE International Symposium on Communications and Information Technology, ISCIT, 2005, pp. 9-12.
11. H. Gao, H. Yang, and X. Wang, “Ant colony optimization based network intrusion feature selection and detection”, In Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 2005, pp. 3871-3875.
12. K. M. Shazzad and J. S. Park, “Optimization of intrusion detection through fast hybrid feature selection”, In Proceedings of the Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT’05), 2005.
13. W. Wong and C. Lai, “Identifying important features for intrusion detection using discriminant analysis and support vector machine”, In Proceedings of the Fifth International Conference on Machine Learning and Cybernetics, Dalian, 2006, pp. 3563-3567.
14. A. Tamilarasan, S. Mukkamala, A. Sung, and K. Yendrapalli, “Feature ranking and selection for intrusion detection using artificial neural networks and statistical methods”, In Proceedings of the International Joint Conference on Neural Networks (IJCNN’06), IEEE, 2006, pp. 4754-4761.
15. A. Zainal, M. A. Maarof, and S. M. Shamduddin, “Feature selection using rough set in intrusion detection”, In Proceedings IEEE TENCON, 2006, pp. 1-4.
16. Y. Li, B. Fang, Y. Chen, and L. Guol, “A lightweight intrusion detection model based on feature selection and maximum entropy model”, International Conference on Communication Technology (ICCT '06), 2006, pp.1-4.
17. S. M. Lee, D. S. Kim, and J. S. Park, “A hybrid approach for real-time

network intrusion detection systems", International Conference on Computational Intelligence and Security, 2007, pp. 712-715.
18. M. J. Fadaeieslam, B. Minaei-Bidgoli, M. Fathy, and M. Soryani, "Comparison of two feature selection methods in intrusion detection systems", Seventh International Conference on Computer and Information Technology, 2007, pp. 83-86.
19. Z. Banković, S. Bojanić, O. Nieto-Taladriz, and A. Badii, "Increasing detection rate of user-to-root attacks using genetic algorithms", The International Conference on Emerging Security Information, Systems and Technologies, 2007, pp. 48-53.
20. Y. Chen, W. Li, and X. Cheng, "Toward building lightweight intrusion detection system through modified RMHC and SVM", 15th International Conference on Networks, ICON 2007, IEEE, 2007, pp. 83-88.
21. C. Tsang, S. Kwong, and H. Wang, "Genetic-fuzzy rule mining approach and evaluation of feature selection techniques for anomaly intrusion detection," *Pattern Recognition*, vol. 40, pp. 2373-2391, 2007.
22. Y. Chen, L. Dai, Y. Li, and X. Cheng, "Building lightweight intrusion detection system based on principal component analysis and C4.5 algorithm", The 9th International Conference on Advanced Communication Technology, ICACT2007, 2007, pp. 2109-2112.
23. M. Panda and M. R. Patra, "Network intrusion detection using Naive Bayes", *International Journal of Computer Science and Network Security*, vol. 7, no. 12, pp. 258–263, 2007.
24. S. Sheen and R. Rajesh, "Network intrusion detection using feature selection and decision tree classifier", IEEE Region 10 Conference, TENCON'08, 2008, pp. 1-4.
25. W. Wang, S. Gombault, and T. Guyet, "Towards fast detecting intrusions: using key attributes of network traffic", The Third International Conference on Internet Monitoring and Protection, 2008, pp. 86-91.
26. M. Bahrololum, E. Salahi, and M. Khaleghi, "Machine learning techniques for feature reduction in intrusion detection systems: a comparison", Fourth International Conference on Computer Sciences and Convergence Information Technology (ICCIT), 2009, pp. 1091-1095.
27. S. Wu and E. Yen, "Data mining-based intrusion detectors", *An International Journal of Expert Systems with Applications*, vol. 36, no. 3, pp. 5605-5612, 2009.
28. H. Tang and Z. Cao, "Machine learning-based intrusion detection algorithms", *Journal of Computational Information Systems*, vol. 5, no. 6, pp. 1825-1831, 2009.
29. S. M. Lee, D. S. Kim, Y. H. Yoon, and J. S. Park, "Quantitative intrusion intensity assessment using important feature selection and proximity metrics", 15th IEEE Pacific Rim International Symposium on Dependable Computing, 2009, pp. 127-134.
30. A. Tajbakhsh, R. Mohammad, and A. Mirzaei, "Intrusion detection using fuzzy association rules", *Applied Soft Computing*, vol. 9, no. 2, pp. 462-469, 2009.
31. L. Xiao, Y. Liu, and L. Xiao, "A Two-step feature selection algorithm adapting to intrusion detection", International Joint Conference on Artificial Intelligence, 2009, pp. 618-622.
32. C. Xiang, T. Yuan, C. Yong-Qin, and Z. Jun-Na, "Robust observation selection for intrusion detection", Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009, pp. 269-272.
33. T. Lee and J. He, "Entropy-based profiling of network traffic for detection of security attack", TENCON, pp. 1-5, 2009.
34. S. Zaman and F. Karray, "Features selection for intrusion detection systems based on support vector machines", 6th IEEE Consumer Communications and Networking Conference (CCNC), 2009, pp. 1-8.
35. K. Khor, C. Ting, and S. Amnuaisuk, "From feature selection to building of bayesian classifiers: a network intrusion detection perspective", *American Journal of Applied Sciences*, vol. 6, no. 11, pp. 1948-1959, 2009.
36. A. Suebsing and N. Hiransakolwong, "Feature selection using euclidean distance and cosine similarity for intrusion detection model", Asian Conference on Intelligent Information and Database Systems (ACIIDS 09), vol. 35, 2009, pp. 86-91.
37. D. Hong and L. Haibo, "A lightweight network intrusion detection model based on feature selection", 15th IEEE Pacific Rim International Symposium on Dependable Computing, 2009, pp. 165-168.
38. M. S. Mok, S. Y. Sohn, and Y. H. Ju, "Random effects logistic regression model for anomaly detection", *Expert Systems with Applications*, vol. 37, no. 10, pp. 7162–7166, 2010.
39. L. M. Ibrahim, "Anomaly network intrusion detection system based on distributed time-delay neural network (DTDNN)", *Journal of Engineering Science and Technology*, vol. 5, no. 4, pp. 457- 471, 2010.
40. H. Nguyen, K. Franke, and S. Petrovic´, "Improving effectiveness of intrusion detection by correlation feature selection", International Conference on Availability, Reliability and Security, 2010, pp. 17-24.
41. T. Chen, X. Pan, Y. Xuan, J. Ma, and J. Jiang, "A naive feature selection method and its application in network intrusion detection", International Conference on Computational Intelligence and Security (CIS), 2010, pp. 416-420.
42. Z. A. Othman, A. A. Bakar, and I. Etubal, "Improving signature detection classification model using features selection based on customized features", 10th International Conference on Intelligent Systems Design and Applications, 2010, pp. 1026-1031.
43. C. Han, Y. Lv, D. Yang, and Y. Hao, "An intrusion detection system based on neural network", International Conference on Mechatronic Science, Electric Engineering and Computer (MEC), IEEE, 2011, pp. 2018-2021.
44. S. Mabu, C. Chen, N. Lu, K. Shimada, and K. Hirasawa, "An intrusion-detection model based on fuzzy class-association-rule mining using genetic network programming", *IEEE Transactions on Systems, Man and Cybernetics—Part C: Applications and Reviews*, vol. 41, no. 1, pp. 130-139, 2011.
45. J. Xian, L. Peiyu, G. Wei, and C. Xuezhi, "An algorithm application in intrusion forensics based on improved information gain", 3rd Symposium on Web Society (SWS), 2011, pp. 100-104.
46. S. Gong, X. Gong, and X. Bi, "Feature selection method for network intrusion based on GQPSO attribute reduction", International Conference on Multimedia Technology (ICMT), 2011, pp. 6365-6368.
47. M. Su, "Real-time anomaly detection systems for denial-of-service attacks by weighted k-nearest-neighbor classifiers", *Expert Systems with Applications*, vol. 38, pp. 3492–3498, 2011.
48. L. Koc, T. A. Mazzuchi, and S. Sarkani, "A network intrusion detection system based on a hidden Naïve Bayes multiclass classifier", *Expert Systems with Applications*, vol. 39, no. 18, pp. 13492-13500, 2012.
49. Y. Li, J. Xia, S. Zhang, J. Yan, X. Ai, and K. Dai, "An efficient intrusion detection system based on support vector machines and gradually feature removal method", *Expert Systems with Applications*, vol. 39, no. 1, pp. 424–430, 2012.
50. A. Osama and Z. A. Othman, "Bees algorithm for feature selection in network anomaly detection", *Journal of applied sciences research*, vol. 8, no. 3, pp. 1748-1756, 2012.
51. Y. Y. Chung and N. Wahid, "A hybrid network intrusion detection system using simplified swarm optimization (SSO)", *Applied soft computing*, vol. 12, no. 9, pp. 3014-3022, 2012.
52. S. S. S. Sindhu, S. Geetha, and A. Kannan, "Decision tree based light weight intrusion detection using a wrapper approach", *Expert Systems with Applications*, vol. 39, no. 1, pp. 129–141, 2012.
53. P. Louvieris, N. Clewley, and X. Liu, "Effects-based feature identification for network intrusion detection", *Neurocomputing*, vol. 121, pp. 265-273, 2013.
54. Z. Karimi, M. M. R. Kashani, and A. Harounabadi, "Feature ranking in intrusion detection dataset using combination of filter methods", *International Journal of Computer Applications*, vol. 78, no. 4, pp. 21-27, 2013.
55. Amrita and P. Ahmed, "A hybrid-based feature selection approach for IDS", In: Meghanathan N., Nagamalai D., Rajasekaran S. (eds) Networks and Communications (NetCom2013). Lecture Notes in Electrical Engineering, Springer, Cham, vol. 284, 2013, pp. 195–211.
56. W. Feng, Q. Zhang, G. Hu, and J. X. Huang, "Mining network data for intrusion detection through combining SVMs with ant colony networks", *Future Generation Computer Systems*, vol. 37, pp. 127-140, 2014.
57. I. Ahmad, M. Hussain, A. Alghamdi, and A. Alelaiwi, "Enhancing SVM performance in intrusion detection using optimal feature subset selection based on genetic principal components", *Neural Computing and Applications*, vol. 24, no. 7-8, pp. 1671–1682, 2014.
58. K. I. Rufai, R. C. Muniyandi, and Z. A. Othman, "Improving bee algorithm based feature selection in intrusion detection system using membrane computing", *Journal of Networks*, vol. 9, no. 3, pp. 523-529, 2014.
59. F. J. Kuang, W. H. Xu, and S. Zhang, "A novel hybrid KPCA

and SVM with GA model for intrusion detection", *Applied Soft Computing*, vol. 18, pp. 178-184, 2014.
60. A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems". *Expert Systems with Applications*, vol. 42, no. 5, pp. 2670-2679, 2015.
61. B. M. Aslahi-Shahri, R. Rahmani, M. Chizari, A. Maralani, M. Eslami, M. J. Golkar, and A. Ebrahimi, "A hybrid method consisting of GA and SVM for intrusion detection system", *Neural Computing and Applications*, vol. 27, no. 6, pp. 1669-676, 2015.
62. X. Z. Wang, "ACO and SVM selection feature weighting of network", *International Journal of Security and Its Applications*, vol. 9, no. 4, pp. 129-270, 2015.
63. H. S. Chae, B. O. Jo, S. H. Choi and T. K. Park, "Feature selection for intrusion detection using NSL-KDD", *Recent Advances in Computer Science*, pp. 960-978, 2015.
64. NSL-KDD dataset for network-based intrusion detection systems. Available on: http://nsl.cs.unb.ca/NSL-KDD/, March 2009.
65. H. Bostani and M. Sheikhan, "Hybrid of binary gravitational search algorithm and mutual information for feature selection in intrusion detection systems", *Soft Computing*, vol. 21, no. 9, pp. 2307–2324, 2015.
66. S. Rastegari, P. Hingston, and C. P. Lam, "Evolving statistical rulesets for network intrusion detection", *Applied soft computing*, vol. 33, pp. 348-359, 2015.
67. S. Ganapathy, P. Vijayakumar, P. Yogesh, and A. Kannan, "An intelligent CRF based feature selection for effective intrusion detection", *The International Arab Journal of Information Technology*, vol. 16, no. 2, pp. 44-50, 2015.
68. M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an intrusion detection system using a filter-based feature selection algorithm", *IEEE Transactions on Computers*, vol. 65, no. 10, pp. 2986–2998, 2016.
69. S. Balakrishnan, K. Venkatalakshmi, and A. Kannan, "A intrusion detection system using feature selection and classification technique", *International journal of computer science and application (IJCSA)*, vol. 3, no. 4 , pp. 145-151, 2016.
70. T. Mehmod and H. B. M. Rais, "Ant colony optimization and feature selection for intrusion detection", *Advances in machine learning and signal processing*, pp. 305-312, 2016.
71. S. Ramakrishnan and S. Devaraju, "Attack's feature selection-based network intrusion detection system using fuzzy control language", *International Journal of Fuzzy Systems*, vol. 19, no. 2, pp. 316–328, 2016.
72. A. I. Madbouly and T. M. Barakat, "Enhanced relevant feature selection model for intrusion detection systems", *International Journal of Intelligent Engineering Informatics*, vol. 4, no. 1, pp. 21-45, 2016.
73. H. Wang, J. Gu, and S. Wang, "An effective intrusion detection framework based on SVM with feature augmentation", *Knowledge-Based Systems*, vol. 136, pp. 130–139, 2017.

AUTHORS PROFILE

Ms. Amrita is an Assistant Professor in Department of Computer Science and Engineering at Sharda University, Greater Noida. She obtained her M. Tech from Bansathali Vidyapith, Rajasthan. She received her M.Tech. in Computer Science from Banasthali Vidyapith, Rajasthan. She is currently pursuing her Ph.D. in Computer Science and Engineering from Sharda University, Greater Noida (U.P.). She has around 19 years of experience in Academics, Software Development Industry and Government Organization.

Shri Kant received Ph.D.(Mathematics) from institute of technology, Deptt of Mathematics, Banaras Hindu University , India. He worked for more than 35 years in Defense Research and Development Organization (DRDO), M/O Defense in various capacities viz: Scientists, Coordinator and Director of a DRDO lab. During this period he has guided a team of scientists working on Pattern Recognition, Cluster Analysis and Soft Computing application in the field of cryptology mainly cryptanalysis. Currently, he is working as a Professor at Research and Technology Development Centre (RTDC), Deptt. of Computer Science and Engineering of Sharda University, India and involved actively in teaching and research mainly in the area of cyber security and medical diagnostics'.

His areas of interest are Special Functions, Cryptology, Pattern Recognition, Cluster Analysis and Data Mining. He has published more than eighty research papers in international and national journals and conferences and also published fifteen technical reports for internal consumption of the departments. He has guided more than thirty PG projects and two PhD theses. He has received commendation certificates and scientist of the Lab award for exhibiting the excellence in pattern recognition application to cryptology.