A HYBRID INTELLIGENT INTRUSION DETECTION SYSTEM USING DOMAIN KNOWLEDGE AND ENSEMBLE LEARNING

Ms. Pranavi Patel
SVIT, Vasad /KJIT, Savali

Ms. Mala Mehta
SVIT, Vasad

ABSTRACT

An intrusion Detection System (IDS) plays a very vital role in protecting systems. For many years researchers have been working on efficient dimensionality reduction procedures, we introduce Domain Knowledge (CIA principles) relevant features to avoid complex methods. Artificial Intelligence has many machine learning algorithms which are effective to detect specific types of attacks only. To overcome this problem we used Ensemble Learning to combine Random Forest, AdaBoost, Naïve Bayes, K-Nearest Neighbor, and Decision Tree. The signature-based IDS and train-test-split method is used to modify and compute. To test our hybrid method we utilized four datasets which are KDDCup99, NSLKDD, UNSW-NB15, and CICIDS, and we gained accuracy of 99%, 96%, 94%, and 99% respectively. Additionally, our method overcomes the False Positive Rate (FPR) problem majorly. Compared to other models our hybrid model shows improved accuracy and a major reduction in FPR.

KEYWORDS: Hybrid-IDS model, Domain Knowledge, CIA Principle, SMOTE, Ensemble Learning, Machine Learning, Intrusion, Datasets and intrusion detection.

1. INTRODUCTION

As almost every field has become digitalized, it has created a need for robust intrusion detection. Machine Learning has become the priority in detecting cyber threats effectively in the last few years. Many researchers have used single algorithms for intrusion detection with different preprocessing methods to improve IDS. However, the methodologies were not effective because every specific algorithm is effective in detecting specific types of attacks only. This creates the need for a combined approach to making IDS effective. Many authors have used different methods for their model. And based on their results, we have figured out some most effective algorithms which are Random Forest (RF), AdaBoost (AB), Naïve Bayes (NB), K-Nearest Neighbor (KNN), and Decision Tree (DT). And to combine their computation efficiency we used the MAX-voting approach of Ensemble Learning (EL). Which is the most used approach for classification problems in EL.

Another factor that affects the performance of IDS is dimensionality reduction. For that, we used Domain Knowledge (DK) features which are CIA principles relevant features. The features are associated with Confidentiality, Integrity, and availability; rules of the network. Many types of studies have been using various complex methods for dimensionality reduction which lead to more computational hazards in IDS. To overcome this issue we have obtained DK features to solve the need for a complex dimensionality reduction procedure in IDS.

We have proposed a hybrid model using both domain knowledge and ensemble learning. As changing natures of the network sometimes old threat patterns are recycled, and also different variants of network feature in change. To ensure that our system gives better performance in different network patterns we

utilized datasets in which two are of early network patterns (KDDCup99, NSLKDD), one of the moderate network patterns (UNSW-NB15) and another of recent network patterns (CICIDS-2017) data sets are used.

In this paper, the next section gives a brief summary of the literature we have studied (Sec. 2). Onward, in Sec.3 a proposed Hybrid model's workflow is defined. Sec.4 includes the results of the proposed IDS model. In Sec. 5 we concluded our study with some future scope.

2. LITERATURE STUDY

In [1] researchers have developed a multi-tree algorithm for combining results of the multiple ML methodologies. Simultaneously they used an adaptive voting method in EL for classification. Additionally, for pre-processing and dimensionality reduction PCA has been utilized. In [2] paper they have utilized efficiency of particle swarm optimization, ant colony algorithm, and genetic algorithm in a hybrid manner to obtain feature selection. Onwards, they have used the two-stage classifier ensemble method to combine all Meta learners and base classifiers. They gained significant improvement in comparison with previously available methods. [3] A novel fuzzy ensemble feature selection with a combined fusion of SVM, KNN, and ANN is obtained. This feature selection method helps to increase accuracy. The authors of [4] developed a Semi-supervised learning model to combine both supervised and unsupervised learning models with EL. They gained an accuracy of 84.54% on KDDCup99.

To better explainability of DK (CIA principles-based features) [5] uses the Black Box-testing method for showing the usability of these features. They have used SMOTE to overcome the imbalance in datasets. Onwards, classification algorithms achieved results that show these features are performing well in detecting unknown attacks, but it does not perform well in detecting known attacks.

The authors of [6] have used the Sparks HDFS system in a real-time environment to overcome issues of Distributed Denial of Service attacks in Ad-hoc networks. An RF classification algorithm was being used and gained a higher detection rate. However, the model has issued more False Positive Rates. [7] Stacked Sparse Auto-encoder is used for feature selection and Support Vector Machine which gives a higher detection rate and also comparatively few False Positives.

The researchers of [8] developed the modified model of a backpropagation neural network by adding LM (Levenberg-Maraurdt) and gaining a significant detection ratio. The authors of [9] have considered authentication problems in IDS. And the proposed model of ANN is a supervised learning model. A considerable improvement is gained inaccuracy (84%) on UNSW-NB15. In [10] neural networks were utilized for MCPS (Medical Cyber-Physical System), which is a new field that requires strong Intrusion detection methods. As for detecting disease from scanned reports, they used the KDDCup99 dataset to test results and gained better detection than previous studies. The model of [11] was designed to improve detection procedures. They considered Gaited Recurrent Units (GRU) for feature selection and RNN with LSTM for classification. To overcome the data imbalance the researchers of [12] invented a model which works on raw data to speed up the classification computation on CNN. The [13] proposed a model which uses a feed-forward deep neural network (FFDNN), to justify the model they used the NSL-KDD dataset. And gained high accuracy in both, binary as well as multi-class classification.

One Side Selection and SMOTE's combined method was used to develop balanced datasets. Then CNN was used to gain spatial features and BiLSTM for temporal features and this hybrid model achieved 80% accuracy [14]. Authors of [15] have tested various machine learning algorithms on different publicly available datasets and concluded that DNN gives a better overall performance on the cost of more timing as DNN is very complex. [16] A Deep Auto-Encoder (DAE) algorithm is being applied on KDDCup99. To avoid overfitting they used the greedy-layer wise fashion DAE and achieved 94.42% (True Positive) and 94.71% (Accuracy). Self-taught learning is being used to feature and dimensionality reduction for computation purposes they used SVM [17, 18].

For dimensionality reduction [19] uses Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) with KNN, SVM, and NB. In [20] weighted SVM with Min-Max Scaler and standardization for dealing with an imbalance in data. AddaBoost shows better performance in detecting cyber threats and also gives a better performance ratio compared to other methods [21]. SVM, ANN, DT, BN, GA, KNN, Fuzzy K-means are more accurate data mining methodologies [22]. SVM, RF, EL's comparative analysis is being performed in [23] and EL outperforms the other two methods.

3. PROPOSED HYBRID IDS MODEL

The proposed hybrid model uses the four benchmarked data sets for cyber security. Due to the various issues in datasets such as data corruption, traffic variety, inconsistencies, ancient contemporary attacks, it has become critical to rely on one dataset's performance for a sustainable IDS model. In ML we have used the supervised learning model. In the supervised learning model, labeled data sets to train the data. Then used that trained data to test the model. For training and testing, we optimized train-test-split methodologies, the portion of training is taken by 70% of the respective datasets and testing has assigned a 30% portion of the respective datasets.

3.1 Domain Knowledge

Domain Knowledge (i.e. CIA Principles) features are on the crucial regulations which are confidentiality, integrity, and availability. We have used DK for signature-based IDS, in previous studies of DK it is majorly known for anomaly detection, however, we have used it for signature-based detection and gained effectively higher accuracy. Different datasets constraints unique features due to that DK features are different respectively. Table 1. Shows DK features of different datasets. Only KDDCup99 and NSL-KDD's features are the same because NSL-KDD is derived from KDDCup99. Although, CICIDS2017's some features are renamed from KDDCup99.

Table: 1 DK Features in Different Datasets

No.	KDDCup99 and NSL-KDD	UNSW-NB15	CICIDS2017
1.	Flow Duration	swin	Ack Flag Count
2.	TotalBackward Packets	dwin	Active Mean

3.	Fwd Packet Length Mean	stcpb	Active Min
4.	Fwd Packet Length Std	dtcpd	Average Packet Size
5.	Flow IAT Mean	smeansz	Bwd IAT Mean
6.	Flow IAT Std	dmeans	Bwd Packet Length Std
7.	Flow IAT Min	trans-depth	Bwd Packets/s
8.	Fwd IAT Mean	res_bdy_len	Fwd IAT Mean
9.	Fwd IAT Min	ct_srv_src	Fwd IAT Min
10.	Bwd IAT Mean	ct_srv_dst	Fwd Packet Length Mean
11.	Fwd PSH Flags	ct_dst_Itm	Fwd packets/s
12.	Fwd Packets/s	ct_src_Itm	Fwd PSH Flags
13.	Bwd Packets/s	ct_dst_sport_Itm	Flow Duration
14.	Syn Flag Count	ct_dst_src_Itm	Flow IAT Mean
15.	PSH Flag Count	-	Flow IAT Min
16.	Ack Flag Count	-	Flow IAT Std
17.	Average Pack Size	-	Init_Win_bytes _Froward
18.	Sub-flow Fwd Bytes	-	PSH Flag Count
19.	Int_Win_bytes_forward	-	Subflow Fwd Bytes
20.	Active Min	-	SYN Flag Count
21.	Idle Mean	-	Total Length of Fwd Packets

3.2 SMOTE

Preprocessing of the datasets plays a very crucial role in every model. Fore mostly, we defined labeled features in binary classification (1 = Intrusions or attacks and 0 = normal traffic). Then to unified datasets for easy utilization of the datasets that must require steps for ML algorithms. So, we have applied StandardScaler. The major issue all the datasets suffer from is a class imbalance in datasets, any one class's majority creates biases in a training model that leads to less accuracy and more false positives. To overcome the imbalance of datasets we used SMOTE (Synthetic Majority Over-sample Technique). SMOTE creates duplicate samples of the minority class to the majority classes samples. Table 2. Shows SMOTEs sample of datasets. CICIDS 2017 is a collection of datasets obtained from working days, in which Monday's datasets do not contain any intrusions that are why it is not included in this study.

Table: 2. Comparative analysis of datasets before and after smote

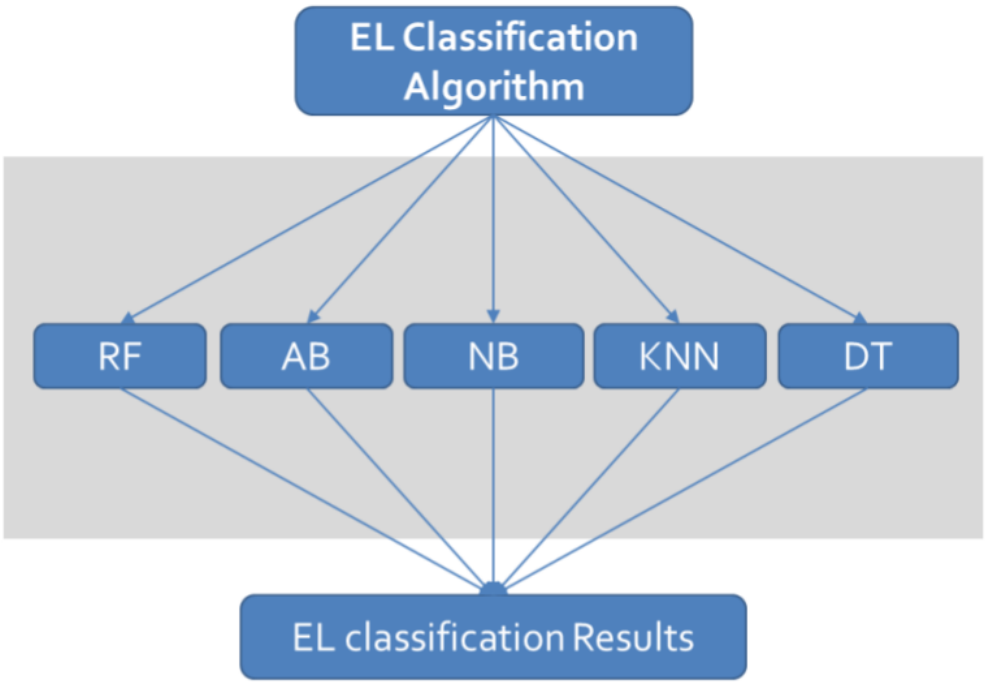
Datasets	Before SMOTE (0)	Before SMOTE (1)	After SMOTE (0)	After SMOTE (1)
KDDCup99	67943	277871	277871	277871
NSL-KDD	47087	41094	47087	47087
UNSW-NB15	25879	31753	31753	31753
Tue CICIDS	9651	302485	302485	302485
Wed CICIDS	308051	176841	308051	308051
Thu morning CICIDS	117752	1504	117752	117752
Thu afternoon CICIDS	201995	26	201995	201995
Fri morning CICIDS	132354	1369	132354	132354
Fri afternoon 1 CICIDS	68214	89807	89807	89807
Fri afternoon 2 CICIDS	89214	111312	111312	111312

3.3 Ensemble Learning

An Ensemble Learning (EL) is used to combine five classification methods which are FR, AB, NB, KNN, and DT respectively. In EL, we used the Max-voting approach for classification. In Max-voting selected models' computation happened individually till they gained classification results. The individual classification results are then combined in EL. The majorly classified class labels become the final classification declarations. Fig 1. Shows the Max-voting approach EL.

RF is a Meta estimator of multiple decision trees that contains various sub-samples of datasets whose average is used to enhance predictive accuracy and control over-fitting. An AB is also a meta-estimator that uses corresponding weights of fitting from various stages, it is also effective in difficult classification natures. NB is a supervised learning algorithm based on the Bayes theorem with naive assumptions of contingent interdependence between pairs of features. KNN is an uncomplicated ML algorithm; its predictions are accurate even though similar patterns are found instead of specified patterns. DT structure is effective with both classification and regression. Its estimation starts from the root to the node for defining class labels.

Figure 1 Max-voting approach of Ensemble Learning



3.4 Hybrid-IDS

Our hybrid model combines DK features with ensemble classification workflow as defined in Fig. 2. The system begins with four standard datasets which are KDDCup99, NSL-KDD, UNSW-NB15, and CICIDS2017. These datasets contain past cybercrimes network patterns. Using those patterns we can make cyberspace a secure environment with sustainable IDS. In DK feature extraction we have removed all the features except mentioned features in Table 1. According to different datasets. StandardScaler converts connections in numeric and normalizes data for ML computations. The imbalance of classes has been overcome using SMOTE oversampling methodologies. Finally, EL combines all the classification results using the Max-voting approach. EL's classification results from the whole procedure become final classification results of connection to be malicious or normal traffic.

4. RESULTS

Accuracy is measured for the basic detection of true positive, false positive, true negative, false negative's detection measures are computer. The recall is portion true positive and false negative. It is mostly used when FN is more sensitive (i.e. medical field). Precision is measured through true positive and ratio of true positive with false-positive. Whenever FP is more important precision is being used (i.e. spam). F1-score is measured based on precision and recall multiplication divided by the summation of both and by multiplying two with results. Tables from 3 to 12 show datasets computations results. Support is the number of records defined by the scores. In this paper, we have taken all the results in between 0.00 to 1.00, where 0.00 denotes 0% and 1.00 denotes 100%.

Figure 2 Proposed Hybrid Model

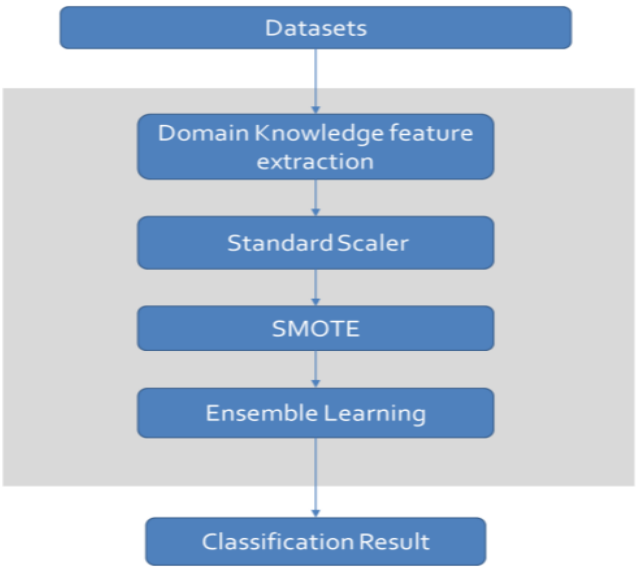


Table 3 computational results of KDDCup99

Measure	Precision	Recall	f-1 score	Support
0	0.98	1.00	0.99	29334
1	1.00	0.99	1.00	118872
Accuracy	-	-	0.99	148206
Macro Average	0.99	1.00	0.99	148206
Weighted Average	1.00	0.99	0.99	148208

We have applied the classic method for Machine Learning evaluation for evaluating results the Precision, Recall, f-1 scores were calculated from Support connection records. Using only accuracy as measures were having the problem of overfitting. In overfitting, algorithms are very effective over trained models, but outside trained records, it does not perform well. To overcome these issues the previous researchers have invented measures Precision, it's used where False Positive is in more concern and Recall, it's used where False Negative in our case both are important because False Positive stops request of genuine connection from accessing services and False Negative let intrusion to successfully run on Networks. For combinational results of Precision and Recall FBeta score is found but the majority of all the research takes Beta's score as 1 and it is called as F-1 Score. Accuracy is a measure of the combinational results of both normal connections and malicious connections so in Precision and Recall it is undefinable. As it's not definable we have denoted it as "-" in tables of computational results.

Table 4 Computational results of NSL-KDD

Measure	Precision	Recall	f-1 score	Support
0	0.95	0.99	0.97	20256
1	0.99	0.93	0.96	17536
Accuracy	-	-	0.96	37792
Macro Average	0.97	0.96	0.96	37792
Weighted Average	0.97	0.96	0.96	37792

Table 5 Computational results of UNSW-NB15

Measure	Precision	Recall	f-1 score	Support
0	0.91	0.98	0.94	11121
1	0.98	0.92	0.95	13579
Accuracy	-	-	0.95	24700
Macro Average	0.95	0.95	0.95	24700

Weighted Average	0.95	0.95	0.95	24700

Table 6 computational results of CICIDS Tuesday

Measure	Precision	Recall	f-1 score	Support
0	1.00	1.00	1.00	129589
1	1.00	1.00	1.00	4184
Accuracy	-	-	1.00	133773
Macro Average	1.00	1.00	1.00	133773
Weighted Average	1.00	1.00	1.00	133773

Table 7 computational results of CICIDS Wednesday

Measure	Precision	Recall	f-1 score	Support
0	1.00	1.00	1.00	131980
1	1.00	1.00	1.00	75831
Accuracy	-	-	1.00	207811
Macro Average	1.00	1.00	1.00	207811
Weighted Average	1.00	1.00	1.00	207811

Table 8 computational results of CICIDS Thursday morning

Measure	Precision	Recall	f-1 score	Support
0	1.00	1.00	1.00	50434
1	0.96	0.99	0.97	676
Accuracy	-	-	1.00	51110
Macro Average	0.98	0.99	0.99	51110
Weighted Average	1.00	1.00	1.00	51110

Table 9 computational results of CICIDS Thursday afternoon

Measure	Precision	Recall	f-1 score	Support
0	1.00	1.00	1.00	86571
1	0.41	0.70	0.52	10
Accuracy	-	-	1.00	86581
Macro Average	0.71	0.85	0.76	86581
Weighted Average	1.00	1.00	1.00	86581

In KDDCup99 we have obtained an accuracy of 0.99 with all the connections which support all the records of data sets which are 148206. Macro and Weighted both averages show 0.99 and 1.00 in both alternatively (Table 3). NSL-KDD has 37792 records total in datasets. The achieved accuracy of the dataset is equivalent to 0.97 (Table 4). For UNSW-NB15 precision according to 0 and 1 is 0.91 and 0.98 respectively (Table 5). Table 6 to Table 12 shows CICIDS's computational results which show CICIDS Tuesday, CICIDS Wednesday, CICIDS Friday afternoon 1, CICIDS Friday afternoon 2 have achieved overall the 1.00 accuracy and also all the measures listed in the table for evaluation have achieved the best results or we can also say ideal system results we have gained. In CICIDS Thursday morning Normal connections are defined correctly with all 0 (normal connections) are achieved 1.00 in support of the 131980. For CICIDS Thursday is also normal connections are defined correctly, however, here the intrusion classification has 0.41, 0.70, and 0.51 Precision, Recall, and f-1 score respectively. The records of CICIDS Friday morning have shown comparatively less Macro Average in Precision which is 0.83. Overall results gained from all the datasets show our method is effective for the classification of malicious behavior and normal connections.

Table 10 computational results of CICIDS Friday morning

Measure	Precision	Recall	f-1 score	Support
0	1.00	0.99	1.00	56713
1	0.67	0.99	0.80	597
Accuracy	-	-	0.99	57310
Macro Average	0.83	0.99	0.90	57310
Weighted Average	1.00	0.99	1.00	57310

Table 11 computational results of CICIDS Friday afternoon 1

Measure	Precision	Recall	f-1 score	Support
0	1.00	1.00	1.00	29504
1	1.00	1.00	1.00	38220
Accuracy	-	-	1.00	67724
Macro Average	1.00	1.00	1.00	67724
Weighted Average	1.00	1.00	1.00	67724

Table 12 computational results of CICIDS Friday afternoon 2

Measure	Precision	Recall	f-1 score	Support
0	1.00	1.00	1.00	38323
1	1.00	1.00	1.00	47618
Accuracy	-	-	1.00	85941
Macro Average	1.00	1.00	1.00	85941
Weighted Average	1.00	1.00	1.00	85941

In KDDCup99 and CICIDS2017 we have reached every score near to perfect in precision, recall, f1-score, and support by using our proposed Hybrid-model. For NSL-KDD every score is being reached to 96% and for UNSW-NB15 it is 95%. Our proposed model's score is giving more accuracy than till now's studies. Additionally Table. 13. Shows comparative analysis of all the classifications results. Hybrid-model not only gives more accuracy but also solves the majority problem of high False Positive Rate (RPR) which is majorly addressed as issues have concurred.

Figure 3 Comparison between only DK and DK + EL

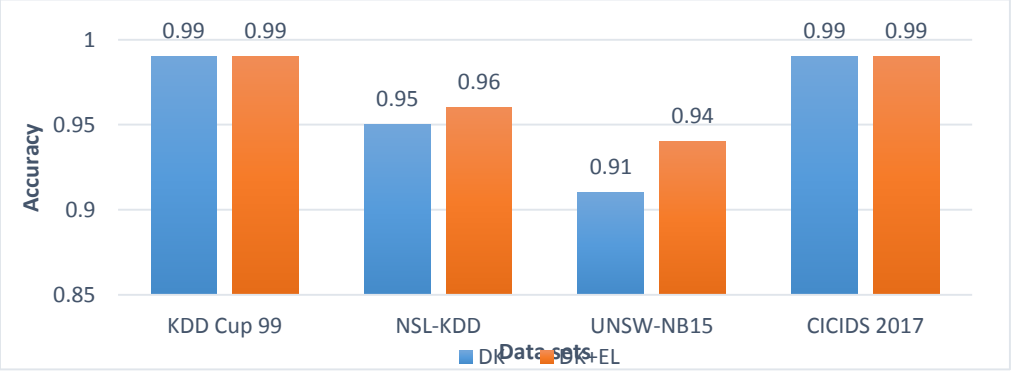


Table: 13. Comparative analysis of all the classifications results.

Dataset	FPR	Recall	Precision	Accuracy	F1-score
KDDCup99	0.02	0.99	1.00	0.99	0.99
NSL-KDD	0.13	0.96	0.97	0.96	0.95
UNSW-NB15	0.19	0.95	0.94	0.95	0.94
CICIDS 2017	0.09	0.99	0.96	0.99	0.96

The accuracy achieved at all different levels of the implementation phase is different; only DK with K-NN for utilized datasets varies from some datasets which show improved accuracy of datasets. Figure 4. Show the comparative results of the accuracy with the Domain Knowledge and Combining it with Ensemble Learning. CICIDS and KDDCup99 do not have any noticeable change in considered decimals points. However, NSL-KDD and UNSW-NB15 show improved results from only applying DK.

The second major concern of developing the system that overcomes False Positive Rate we effectively have reduced the FPR in our proposed hybrid model. Figure 4. Shows the datasets relevant FPR. We significantly gained 0.02 in KDDCup99. And UNSW-NB15 shows relevantly more than other datasets recorded FPR. NSL-KDD and CICIDS2017 have 0.13 and 0.09 respectively. Figure 5. Shows a Radar chart of over-evaluation of our proposed model according to the respective datasets.

Figure 4 Comparison between only DK and EL

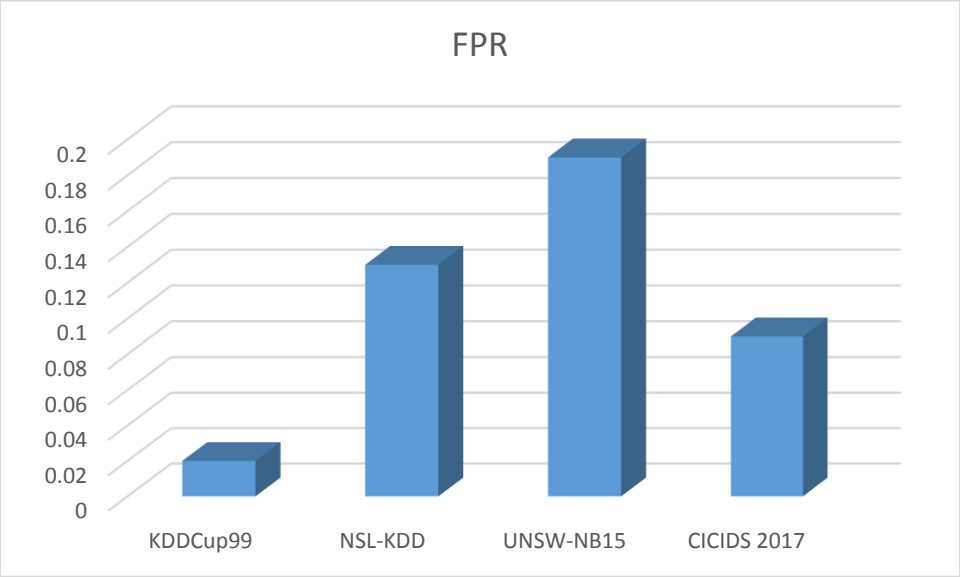
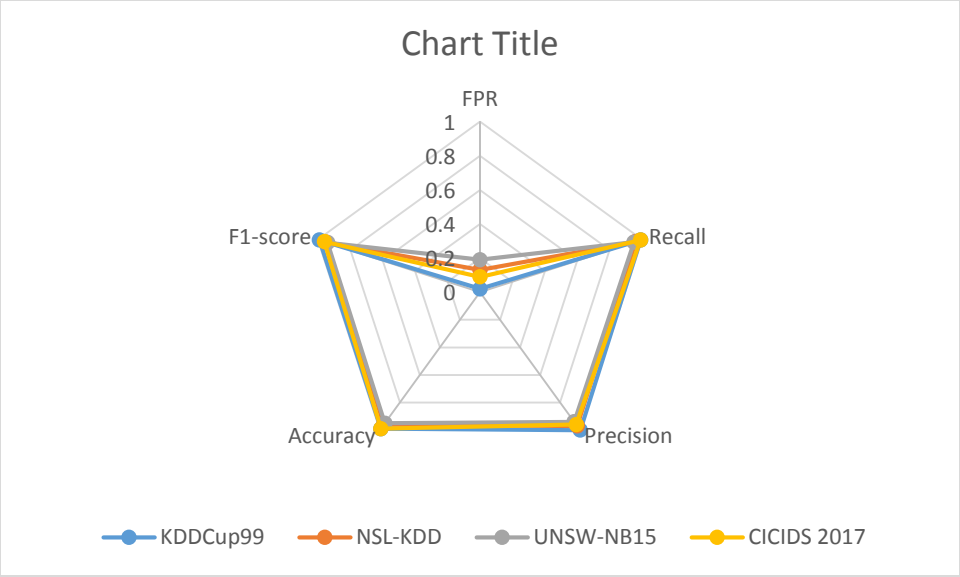


Figure 5 Model Evaluation Graph



4. CONCLUSION

Due to the fluctuating nature of Intrusions and also network patterns, it is required to test models on different kinds of network patterns with different malware records in unique benchmarked datasets. Currently, many IA techniques are being used for feature selection. However, the majority of them suffer from higher False Positive Rates (FPR) in our studies we have been concerned with all the measures for Intrusion detection. As intruders have advanced in attacking, we also need to make our system free from different vulnerabilities. Sustainable IDS has become a must while the majority of fields have been shifted to online. Our developed model shows sustainability over recorded Intrusions from previous years.

In future work we would like to work on two measures effectively, the initial one is even reducing the false positive rate and the final one is performance evaluation on a real-time network rather than recorded sets.

REFERENCES

[1]. Xianwei Gao, Chun Shan, Changzhen Hu, Zequn Niu, And Zhen Li; An Adaptive Ensemble Machine Learning Model For Intrusion Detection; Volume 7, (2019) IEEEAccess.
[2]. Bayu Adhi Tama 1, Marco Comuzzi1, And Kyung-Hyune Rhee; Tse-Ids: A Two-Stage Classifier Ensemble For Intelligent Anomaly-Based Intrusion Detection System; Volume 7, (2019); IEEEAccess.
[3]. Pullagura Indira Priyadarsini1 , G. Anuradha2; A Novel Ensemble Modeling For Intrusion Detection System; Volume 10; (2019); IJECE
[4]. Kehe Wu, Zuge Chen, And Wei Li; A Novel Intrusion Detection Model For A Massive Network Using Convolutional Neural Networks; Volume 6, (2018); IEEEAccess.
[4]. Sheikh Rabiul Islam, William Eberle, Sheikh K. Ghafoor, Ambareen Siraj, Mike Roger; Domain Knowledge Aided Explainable Artificial Intelligence For Intrusion Detection And Response; Arxiv:1911.09853v2 [Cs.Ai] 22 Feb (2020); AAAI.
[5]. Ying Gao, Hongrui Wu, Binjie Song, Yaqia Jin, Xiongwen Luo, And Xing Zeng; A Distributed Network Intrusion Detection System For Distributed Denial Of Service Attacks In Vehicular Ad Hoc Network; Volume 7, (2019); IEEEAccess.

[6]. Binghao Yan And Guodong Han; Effective Feature Extraction Via Stacked Sparse Autoencoder To Improve Intrusion Detection System; Volume 6, (2018); IEEEAccess.
[7]. Aimin Yang 1,2, Yunxi Zhuansun1,2, Chenshuai Liu2, Jie Li 2,3, And Chunying Zhang1,2; Design Of Intrusion Detection System For Internet Of Things Based On Improved Bp Neural Network; Volume 7, (2019); IEEEAccess.
[8]. Sohaib Hanif, Tuba Ilyas, Muhammad Zeeshan; Intrusion Detection In Iot Using Artificial Neural Networks On Unsw-15 Dataset; February (2020); ResearchGate.
[9]. Nishat Mowla, Inshil Doh, Kijoon Chae; Evolving Neural Network Intrusion Detection System For Mcps; Volume 6, (2017), TACT.
[10]. Congyuan Xu, (Student Member, Ieee), Jizhong Shen, Xin Du, And Fan Zhang, (Member, Ieee); An Intrusion Detection System Using A Deep Neural Network With Gated Recurrent Units; Volume 6, (2018); IEEEAccess.
[11]. Kehe Wu, Zuge Chen, And Wei Li; A Novel Intrusion Detection Model For A Massive Network Using Convolutional Neural Networks; Volume 6, (2018); IEEEAccess.
[12]. Kaiyuan Jiang, Wenya Wang, Aili Wang, And Haibin Wu; Network Intrusion Detection Combined Hybrid Sampling With Deep Hierarchical Network; Volume 8, (2020); IEEEAccess.
[13]. Kaiyuan Jiang, Wenya Wang, Aili Wang, And Haibin Wu; Network Intrusion Detection Combined Hybrid Sampling With Deep Hierarchical Network; Volume 8, (2020); IEEEAccess.
[14]. R. Vinayakumar 1, Mamoun Alazab2, (Senior Member, Ieee), K. P. Soman1, Prabaharan Poornachandran3, Ameer Al-Nemrat4, And Sitalakshmi Venkatraman5; Deep Learning Approach For Intelligent Intrusion Detection System; Volume 7, (2019); IEEEAccess.
[15]. Fahimeh Farahnakian, Jukka Heikkonen; A Deep Auto-Encoder Based Approach For Intrusion Detection System; February 11 ~ 14, (2018); ICACT.
[16]. Majjed Al-Qatf, Yu Lasheng, Mohammed Al-Habib, And Kamal Al-Sabahi; Deep Learning Approach Combining Sparse Autoencoder With Svm For Network Intrusion Detection; Volume 6, (2018); IEEEAccess.
[17]. Admir Midzic1 , Zikrija Avdagic2, And Samir Omanovic3; Intrusion Detection System Modeling Based On Learning From Network Traffic Data; Volume 12, (2018); KSII.
[18]. Husam Ibrahim Alsaadi1,2, Rafah M. Almuttairi3, Oguz Bayat1, And Osman Nuri Ucani1; Computational Intelligence Algorithms To Handle Dimensionality Reduction For Enhancing Intrusion Detection System; 293-308 (2020); Journal Of Information Science And Engineering
[19]. Alaeddin Alabdallah1, Mohammed Awad2; Using Weighted Support Vector Machine To Address The Imbalanced Classes Problem Of Intrusion Detection System; Volume 12, (2018); KSII.
[20]. D. Sudaroli Vijayakumar1 & S. Ganapathy2; Machine Learning Approach To Combat False Alarms In Wireless Intrusion Detection System; Volume 11, (2018); Computer And Information Science
[21]. Fadi Salo 1, Mohammadnoor Injadat 1, Ali Bou Nassif 1,2, Abdallah Shami 1, And Aleksander Essex 1; Data Mining Techniques In Intrusion Detection Systems: A Systematic Literature Review; Volume 6, (2018); IEEEAccess.
[22]. Iftikhar Ahmad 1, Mohammad Basheri1, Muhammad Javed Iqbal2, And Aneel Rahim3; Performance Comparison Of Support Vector Machine, Random Forest, And Extreme Learning Machine For Intrusion Detection; Volume 6, (2018); IEEEAccess.