Multimedia Tools and Applications (2021) 80:29643–29656
https://doi.org/10.1007/s11042-021-11100-x

Intrusion detection by machine learning for multimedia platform

Chih-Yu Hsu[1] · Shuai Wang[2] · Yu Qiao[3]

Received: 5 July 2020 / Revised: 27 April 2021 / Accepted: 21 May 2021 /
Published online: 7 July 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

The multimedia service company, Netflix, increased the number of new subscribers during the Coronavirus pandemic age. Intrusion detection systems for multimedia platforms can prevent the platform from network attacks. An intelligent intrusion detection system is proposed for the security IP Multimedia Subsystem (IMS) based on machine learning technology. For increasing the accuracy of the classifiers, it is vital to select the critical features to construct the intrusion detection system. Two-class classifiers, including the Decision Tree, Support Vector Machine, and Naive Bayesian, are selected to evaluate intrusion detection accuracy. According to the three classifiers' accuracy values, the most critical features are selected based on the features' ranking orders. Six critical features are selected:Service, dst_host_same_srv_rate, Flag, Protocol Type, Dst_host_rerror_rate, and Count. Numerical comparison with state_of_the_art shows that critical features improve intrusion detection accuracy, which can be better than the deep learning method.

Keywords Intrusion detection · Support vector machine · Decision tree · Naive Bayesian classifier · Machine learning · Streaming service · Coronavirus pandemic

✉ Yu Qiao
amberjoe1214@163.com

Chih-Yu Hsu
61201903@fjut.edu.cn

Shuai Wang
1196154403@qq.com

[1] Fujian Provincial Key Laboratory of Big Data Mining and Applications, School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, China

[2] School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, China

[3] STEM, University of South Austrilia, Mawson Lakes, SA 5095, Australia

1 Introduction

Due to the Coronavirus pandemic, the streaming service such as Netflix has increased in the first quarter of 2020. Intrusion detection system for multimedia platforms can prevent the platform from being attacked. Modern communication infrastructure that includes the IP Multimedia Subsystem (IMS) and Voice over IP (VoIP) are suffering by attacks and unknown threaten [2]. The multimedia platform provides services to users through the Internet. With the advancement of technologies, cyber-attacks are renewing rapidly, and businesses have a higher risk of information security. A multimedia traffic classification scheme for intrusion detection systems is important [24]. Malicious Intrusion in the network is increasing every day. They are deliberately unauthorized, illegal attempts to access, manipulate, or take possession of an information system Network to render them unreliable or unusable. Intrusion Detection is the process of identifying various events occurring in a system/network and analyzing them for the possible presence of Intrusion and responding to malicious activities. Now Intrusion Detection has become the priority and an important task of Information Security administrators. A system deployed in a network is vulnerable to various attacks and needs to be protected against attacks [1]. Intrusion Detection System (IDS) was originally a software application. However, considering the real-time requirements, special equipment was developed to monitor the network for malicious activities or violations of policies. The management system of these devices collects any malicious activities, violations, security information and event logs and aggregates them into reports. Some IDS can respond when an intrusion is detected, so these are classified as intrusion prevention systems (IPS). The Intrusion Detection System (IDS) is designed to be monitoring and analyzing network traffic and events in the system to discover unauthorized access to computers in the network and play a vital role in protecting the organization's security. The aim of an IDS is to protect the system from unauthorized access, so it collects information about a given network environment, removes redundant information, and makes decisions based on whether the activity is normal or intrusive. Researchers have used various approaches such as data mining, soft computing, Machine Learning, Statistical Techniques, Bayesian Techniques, Artificial Neural Networks, and Evolutionary Computing. Network Anomaly Detection has achieved the purpose of making decisions based on whether the activity is normal or intrusive. Machine learning [11, 32, 36] has some popular classifiers such as Naive Bayesian classifier, Decision Tree, and Support Vector Machine(SVM).Naive Bayesian classifiers uses the smallest classification error probability or the lowest average risk with a given cost. The SVM classifier is a supervised learning algorithm that classifies the data in binary form. Although we can find some survey of cloud-based network intrusion detection analysis, but the intrusion detection research by machine learning applied for multimedia platform is hardly found on the internet. The paper is focus on the apply the technology on multimedia platform.

Supervised learning is given a bunch of samples with attributes and categories. These categories are determined in advance. Then the classifier obtained through learning can give the correct classification to the newly appeared objects. SVM was proposed in 1964 [4]. The decision boundary of SVM is the maximum margin hyperplane that solves the learning samples [7]. Since the 1990s, it has been rapidly developed and derived a series of improved and extended algorithms. SVM has been used in pattern recognition problems such as facial recognition [3] and text classification [15]. In machine learning, the decision tree is a predictive model representing a mapping relationship between object attributes and object values [14]. In the late 1980s and early 1990s, the researcher I. Ross. Quinlan developed a binary classification tree algorithm called ID3, and Quinlan later proposed C4.5 [34, 35]. Machine

learning algorithms for training classifiers have difficulty reducing the number of features, and it is an exciting challenge for researchers. To use machine learning algorithms effectively, preprocessing of the data is essential. Feature selection is one of the most frequent and vital data preprocessing techniques and has become an indispensable component of the machine learning process [9, 20, 26]. Feature selection is the process of selecting relevant features or a candidate subset of features. The evaluation criteria are used to getting an optimal feature subset. To find the high dimensional data of the optimal feature subset is a difficult task [21]. In general, feature selection refers to applying statistical tests to inputs, given a specified output, to determine which columns are more predictive of the output. The algorithms used in measuring the importance of the features include statistic methods, Pearsons's or Kendall's correlation, mutual information scores and chi-squared values. In this paper, the research focus on finding the best features applied to three classifiers for IDS. The features are ranked by predictive power, and the best features are selected based on their scores for defined metrics. The contribution of this paper is to propose a management system of Intrusion Prevention System (IPS) apply the technology on multimedia platform. The IPS including IDS helps in monitoring of all regular and normal patterns of traffic and sends alerts in case of any kind of deviation from the normal pattern.

2 Intrusion detection system for multimedia platform

IP Multimedia Subsystem (IMS) is a multimedia platform developed to provide distinct network services like voice, data, and video. The idea of IMS [30] is to assimilate voice communication and Internet technologies. It includes the sets of core network functional entities and interfaces used by service providers to provide services based on the Session Initiation Protocol (SIP) [33]. IMS promises to provide multi-services, miscellaneous access networks, IP based secure, and reliable network.

2.1 SIP flooding attack

The architecture of attack detection of the SIP flooding attack is shown in Fig. 1. Intrusion Detection System (IDS) can be implemented by software such as OpenIMSCore to achieve the network's security by observing the abnormal behavior of the network [10]. The open-source IMS client and the OpenIMSCore can probe the packets in the traffic network. The IDS is placed on workstations then it is known as Host-based IDS.

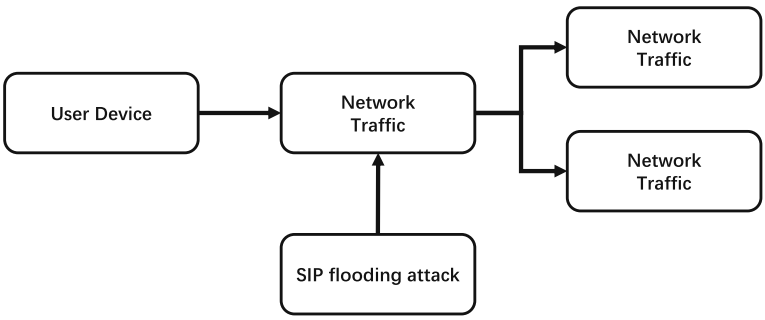

Fig. 1 Architecture of SIP flooding attack detection

2.2 The proposed intrusion prevent system

The Intrusion Prevention System (IPS) [25, 29, 31, 39] is extended from Intrusion Detection System (IDS), as shown in Fig. 2(a), helps in monitoring of all regular and normal patterns of traffic and sends alerts in case of any kind of deviation from the normal pattern. In our designed IPS, it can learn to protect intrusion through the log file when the attack events happen as shown in Fig. 2(b). Because people can be allowed access over the network, the packets sent from intruders are merged into the network traffic as input. It is necessary to monitor the all incoming and outgoing traffic. Intelligent intrusion detection has a trained model for deciding whether the behavior of network traffic is normal or not. If the trained model detects the abnormal behavior category, it is identified as the attack and raised the alarm as a signal to the respected devices' owner. Five labels of the data packets are normal, Denial of Service Attack (DoS), User to Root Attack(U2R), Remote to Local Attack(R2L) that can be probed and stored in the database.

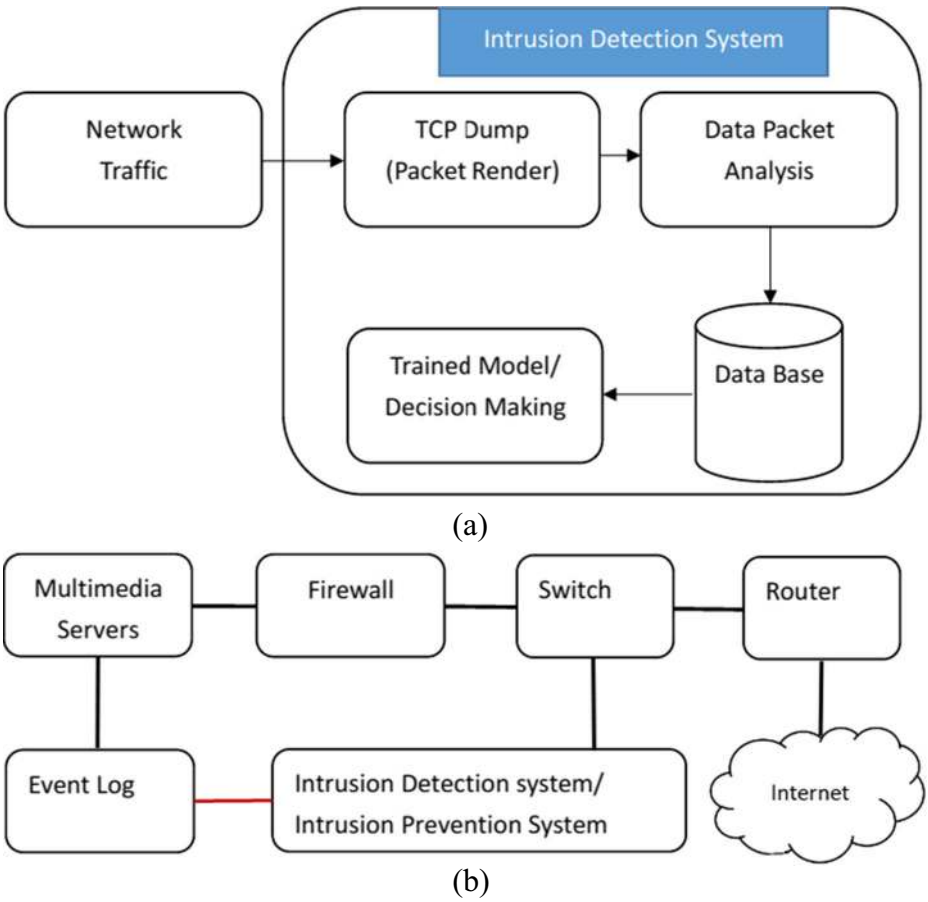

Fig. 2 Intelligent intrusion detection system and intrusion prevention systems

2.3 Machine learning technologies

Machine Learning (ML) is used to analyze and construct the system based on the data sets [5, 13, 22, 37, 38]. There are mainly three types of learning techniques based on labeled data, i.e., Supervised, unsupervised, and semi-supervised learning. Common machine learning algorithms are Support Vector Machine (SVM) [8, 23, 28]. Naïve-Bayes classifier [12, 27], K-nearest neighbor (KNN) [17, 40], artificial neural network (ANN) [16, 18, 19], deep neural network (DNN) [6], and so on. Figure 3 shows how to select the best features by permutation feature importance. Permutation feature importance measures the increase in the prediction error of the model after we permuted the feature's values, which breaks the relationship between the feature and the true outcome. The model training procedure includes input training data, feature extraction, training model, evaluation and validation. The trained model can be used to test the new input data and make decision whether the packet traffic is normal.

3 Theory of classification

This section explains the theory of the classifiers such as Decision Tree (DT), Support Vector Machine (SVM), and Naïve Bayes (NB).

3.1 Binary classification tree

A procedure for growing a binary classification tree (BCT) is described. A space R^d is a d-dimensional cubic surface that contains the training data points $x_i = (x_{i1}, x_{i2}, \ldots, x_{id}), i = 1, \ldots, n$ A plane in R^d splits a region R_{k-1} into two subregions R_k and R'_k(k >= 1). The Function $E(a)$ is to calculate fraction of points in $x_i \in R$ missclassified by a majority in region R. The plane has two parameters j and s, and the optimal j and s minimize equation is $E\left(R_k(j,s)\right) + E\left(R'_k(j,s)\right)$. The function E is as follow:

$$E(R) = \begin{cases} \frac{N_0}{N_R}, \text{if } N_0 < N_R \\ \\ \frac{N_1}{N_K}, \text{if } N_1 < N_0 \end{cases} \tag{1}$$

where $\mathrm{R_k(j,s)} = \left\{\mathrm{x_i} \in \mathrm{R_k} \mid \mathrm{x_{ij}} > \mathrm{s}, 1 < \mathrm{j} < \mathrm{d}\right\}$, $\mathrm{R'_k(j,s)} = \left\{\mathrm{x_i} \in \mathrm{R_k} \mid \mathrm{x_{ij}} \leq \mathrm{s}, 1 < \mathrm{j} < \mathrm{d}\right\}$. The N_0 is the number of points x_i with label 0. The N_1 is the number of points x_i with

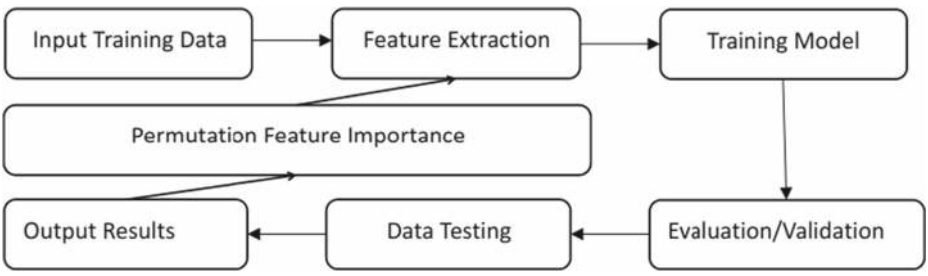

Fig. 3 Model training and testing procedure

label 1. The N_R Is the number of points and $N_R = N_0 + N_1$. Stopping criteria is only $x_i \in R$ one points in R_k.

3.2 Support vector machine classifier

Given a training data set D has n points and corresponding label set L.$\vec{x} = (x_{i1}, x_{i2}, \ldots, x_{id})$ is a vector in a d-dimensional space $\mathrm{R^d}$ and $x_i \in D$. y_i is a lable and $y_i \in \mathrm{L} = \{1, -1\}$. It is an optimization problem to find the "maximum-margin hyperplane" for the SVM classifier. The hyperplane is the set of points $\vec{x} \in R^d$ satisfying $\vec{w} \cdot \vec{x} - b = 0$. where $\vec{w}$ is the normal vector to the hyperplane and the parameter $\frac{\mathrm{b}}{\|\vec{w}\|}$ determines the offset of the hyperplane from the origin along the normal vector $\vec{w}$. Two parallel hyperplane that separate two classes of data are the region bonded by these two hyperplanes in called the "margin", and the maximum-margin hyperplane is the middle of these two hyperplanes as follows:

$$\vec{\mathbf{w}} \cdot \vec{\mathbf{x}} - b = \begin{cases} +1, & \text{for } y_i = 1 \\ \\ -1, & \text{for } y_i = -1 \end{cases} \tag{2}$$

Where $\vec{\mathbf{x}}$ denotes a input feature vector. The distance between these two hyperplane is $\frac{b}{||\vec{w}||}$, so to maximize the distance equal to minimize $||\vec{w}||$. The optimization problem is formulated as follow:

$$\min_{\vec{w},b} ||\vec{\mathrm{w}}|| \tag{3}$$

In (3), it subject to $y_i \cdot (\vec{\mathbf{w}} \cdot \vec{\mathbf{x}} - b) \geq 1$, for $i = 1, 2, \ldots, n$

3.3 Naïve Bayesian classifier

The Naive Bayes algorithm is based on Bayes' theorem. The formula of Bayes' theorem is as follows:

$$P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)} \tag{4}$$

$P(A)$ is a priori probability. $P(B \mid A)$ is the conditional probability. $P(A \mid B)$ is the posterior probability. In addition, $P(B)$ is the probability of B occurring under different given conditions. It can be expressed as (5):

$$P(B) = P(B \mid A_1) P(A_1) + P(B \mid A_2) P(A_2) + \ldots + P(B \mid A_n) P \quad (A_n) \tag{5}$$

$$D = \left\{ \left(x^{(1)}, y^{(1)} \right), \left(x^{(2)}, y^{(2)} \right), \ldots, \left(x^{(N)}, y^{(N)} \right) \right\} \tag{6}$$

There are N data in D. Each data has n characteristics. y is the class corresponding to x. There are k classes. The method of determining the class to which a given x belongs is as follows. For a given x, by Bayes' theorem:

$$P(C_k \mid x) = \frac{P(x \mid C_k) P(C_k)}{P(x)} \tag{7}$$

For the naive Bayes classifier, it is assumed that n features are independent of each other. Then bring (8) into (7) and bring $P(x)$ in (7) into the full probability formula. Then formula 9 is the naive Bayes model:

$$P(x \mid C_k) = P(x_1, x_2, \ldots, x_m \mid C_K) = \prod_{i=1}^{n} P(x_i \mid C_k) \tag{8}$$

	Actually: Positive	Actually: Negative
Predicted: Positive	TP	FP
Predicted: Negative	FN	TN

Fig. 4 Confusion matrix

$$P\left(C_k \mid x\right) = \frac{P\left(C_k\right)\prod_{i=1}^{n} P\left(x_i \mid C_k\right)}{\sum_{k=1}^{k}\left[P\left(C_k\right)\prod_{i=1}^{n} P\left(x_i \mid C_k\right)\right]} \tag{9}$$

The symbols in Fig. 4 are explained as follows. True Positive means we predicted the object as positive and it is actually positive. True Negative means that we predicted the object is negative and it is actually negative. False Positive means we predicted the object as positive but it is actually negative. False Negative means we predicted the object as negative but it is actually positive.

4 Experimental results and discussion

The database used in this paper is the NSL-KDD test dataset, and each sample in the data set has 41 features. There are five labels of the data are normal, Denial of Service Attack (DoS), User to Root Attack(U2R), Remote to Local Attack(R2L), and probe. All the labels except the normal indicate the different attacks in the dataset. The NSL-KDD data set is divided into two categories-normal and attack. The main challenge of the intrusion detection model is to achieve maximum accuracy with a minimum false alarm rate (FP). The results are shown in three subsections. In the Sections 3.1, the decision tree classifier model is used for testing and finding the optimal features. In the Section 3.2, the SVM classifier is used for testing and finding the optimal features. In the Sections 3.3, the Bayes classifier is used for testing and finding the optimal features.

Table 1 Parameters of decision tree classifier

Parameters	Value
Maximum number of Leaves	20
Minimum Leaf Instances	10
Learning Rate	0.2
Number of Levels	100
Allow Unknown Levels	True

Table 2 Confusion Matrix of decision tree classifier

N = 9018	Actual Positive	Actual Negative
Predicted Positive	3889	40
Predicted Negative	56	5033

4.1 Decision tree classifier

Two-Class Boosted Decision Tree is a binary classifier used for testing and finding the optimal features. Table 1 shows the parameter values of the decision tree. The threshold value is set to 0.5. The Confusion matrix is as shown in Table 2. In the fusion matrix, the value of True Positive is 3889, the value of False Positive is 40, the value of False Negative is 56, and the value of True Negative is 5033. The Accuracy of the decision tree classifier reaches 98.9%, and the corresponding area under the curve (AUC) value reaches 0.999. The resulting receiver operating characteristic (ROC) curve is shown in Fig. 5. The permutation feature importance scores the features in the decision tree model, as shown in Table 3. The first column is the rank order, and the second column is the feature name. The third column is the score of the feature importance. The top 10 records are listed in rank order.

4.2 Support vector machine classifier

The Support Vector Machine (SVM) classifier is a binary classifier. Table 4 shows the parameter values of the SVM. The threshold is 0.5. Table 5 is the Confusion Matrix of SVM. The value of True Positive is 3711. The value of False Positive is 231. The value of False Negative is 234 and the value of True Negative is 4842. The Accuracy of the decision tree classifier reaches 94.8%, and the corresponding area under the curve (AUC)

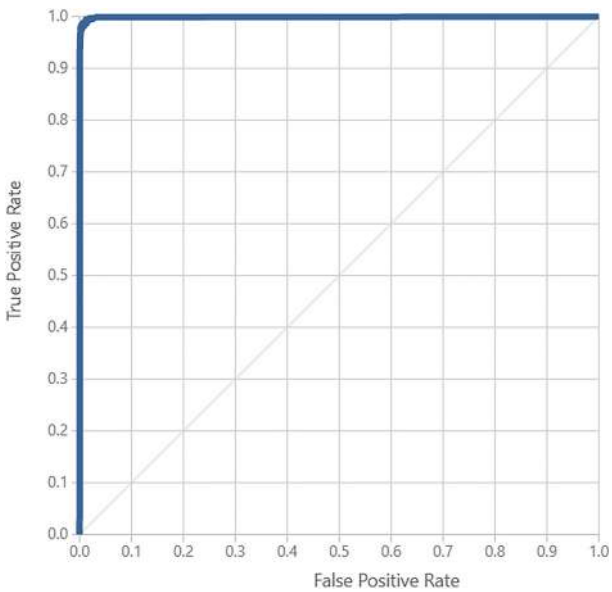

Fig. 5 The ROC curve of the decision tree classifier

Table 3 Feature score of the decision tree classifier

Rank	Feature	Score
1	src_bytes	0.145154
2	Service	0.040696
3	srv_count	0.008694
4	dst_host_same_src_port_rate	0.005323
5	dst_bytes	0.004768
6	Duration	0.004214
7	Protocol_type	0.002994
8	Count	0.00255
9	Flag	0.002107
10	dst_host_error_rate	0.001442

Table 4 Parameters of SVM classifier

Parameters	Value
number of iterations	1
Lambda	0.001

Table 5 Confusion Matrix of SVM classifier

N = 9018	Actual positive	Actual negative
Predicted Positive	3711	231
Predicted Negative	234	4842

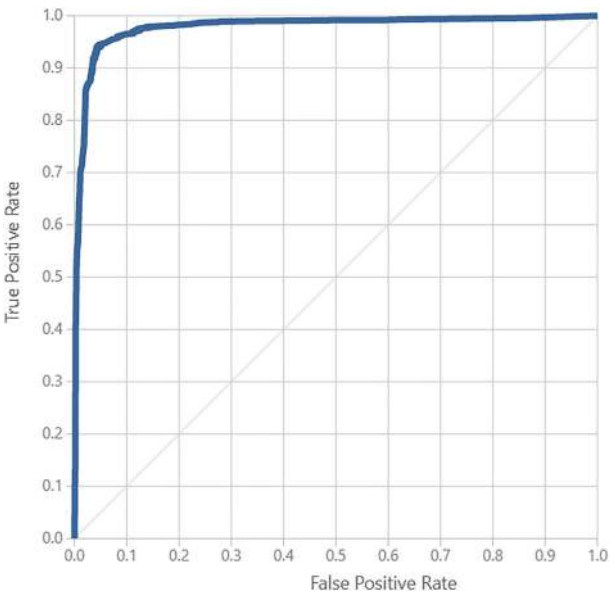

Fig. 6 The ROC curve of the SVM classifier

Table 6 The feature score of the SVM classifier

Rank	Feature	Score
1	Service	0.264804
2	dst_host_rerror_rate	0.065646
3	dst_host_same_srv_rate	0.055777
4	Flag	0.020958
5	dst_host_srv_count	0.016301
6	same_srv_count	0.011865
7	Protocol_type	0.011311
8	dst_host_count	0.00998
9	diff_srv_rate	0.009093
10	Count	0.0054434

Table 7 Parameters of Naïve Bayesian classifier

Parameters	Value
Allow Unknown Levels	True
Random Number Seed	2342
Training Iteration Count	30
Add Bias	True

Table 8 Confusion Matrix of Naïve Bayesian classifier

N = 9018	Actual Positive	Actual Negative
Predicted Positive	3695	246
Predicted Negative	250	4827

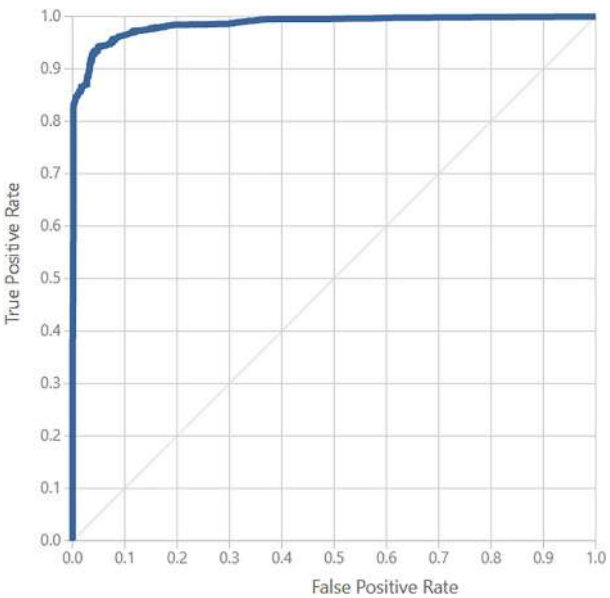

Fig. 7 The ROC curve of the Naïve Bayesian classifier

Table 9 The feature score of the Naïve Bayesian classifier

Rank	Feature	Score
1	Service	0.211244
2	Count	0.064538
3	Flag	0.055666
4	Protocol_type	0.05367
5	dst_host_same_srv_rate	0.041476
6	srv_count	0.03992
7	dst_host_rerror_rate	0.031825
8	same_srv_rate	0.005101
9	srv_rerror_rate	0.003327
10	rerror_rate	0.002772

value reaches 0.985. The resulting receiver operating characteristic (ROC) curve is shown in Fig. 6. The permutation feature importance scores the features in the SVM classifier, as shown in Table 6. The first column is the rank order, and the second column is the feature name. The third column is the score of the feature importance. The top 10 records are listed in rank order.

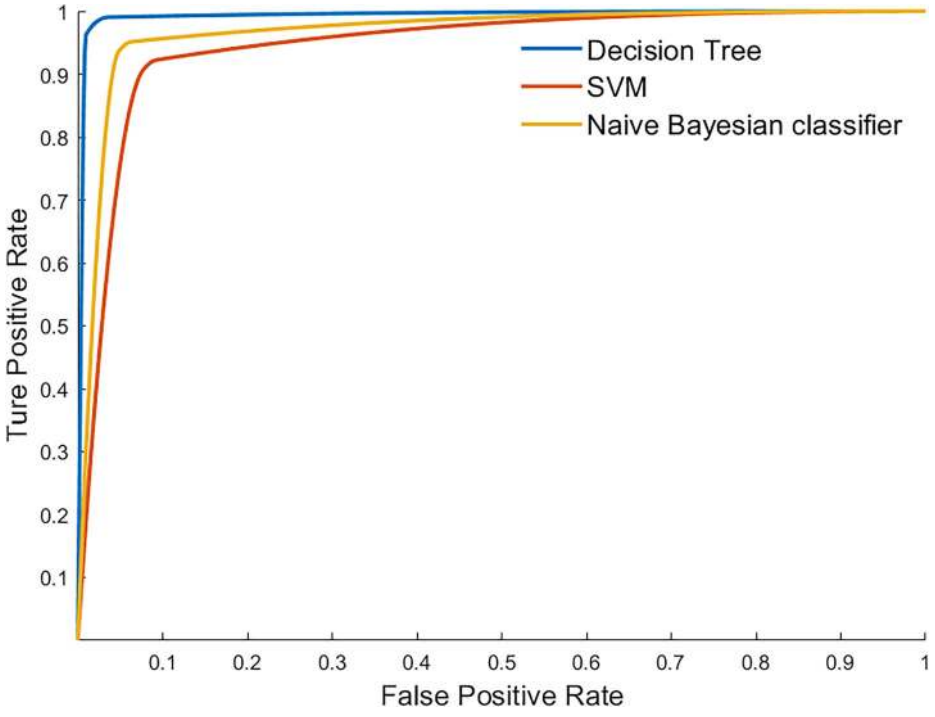

Fig. 8 The ROC curves of the Decision Tree, SVM and Naïve Bayesian classifier

Table 10 Performance comparison between classifiers and state-of-the-art

	Accuracy	Precision	Recall	F1 Score
Decision Tree	0.989	0.990	0.986	0.988
SVM	0.948	0.901	0.908	0.904
Naïve Bayesian classifier	0.944	0.938	0.937	0.937
Deep Learning [6]	0.915	0.930	0.921	0.928

4.3 Naïve Bayesian classifier

The Naïve Bayesian classifier is a binary classifier. Table 7 shows the parameter values of the Naïve Bayesian. The threshold is 0.5. Table 8 is the Confusion Matrix of Naïve Bayesian classifier. The value of True Positive is 3695. The value of False Positive is 246. The value of False Negative is 250 and the value of True Negative is 4827. The Accuracy of the Naïve Bayesian classifier reaches 94.4%, and the corresponding area under the curve (AUC) value reaches 0.978. The resulting receiver operating characteristic (ROC) curve is shown in Fig. 7. The permutation feature importance scores the features in the Naïve Bayesian classifier as shown in Table 9. The first column is the rank order, and the second column is the feature name. The third column is the score of the feature importance. The top 10 records are listed in rank order.

4.4 Comparison

Machine Learning algorithms are developing on the rising situation. Every year new techniques are presented that update the current leading algorithms. It is hard to define state of art since there is not certain algorithm capable of solving all kind of ML problems. The need of ML algorithms really varies with the constraints of the tasks. However in some sense we might list the well performing algorithms in their suitable use-cases with the best result such as SVM, Decision Tree, Naïve Bayesian classifier. For the intrusion detection task can be presented the comparisons between classifiers and deep learning model [6]. Figure 8 shows the ROC curves of the Decision Tree, SVM and Naïve Bayesian classifier. Their area under the curve (AUC) individually are 0.999, 0.985 and 0.978. Performance comparisons between classifiers and state-of-the-art methods are listed in the Table 10. The Decision Tree classifier has the best performance.

5 Conclusion

An intelligent intrusion detection system is proposed for the security IP Multimedia Subsystem (IMS) based on machine learning technology. For increasing the accuracy of the classifiers, it is vital to select the critical features to construct the intrusion detection system. The decision tree, SVM, and Bayesian are binary classifiers that are used to test the NSL-KDD data set. Based on the experimental results, six critical features affecting the accuracy are respectively “Service”, “dst_host_same_srv_rate”, “Flag”, “Protocol_type”, “Dst_host_rerror_rate” and “Count”. The values of intrusion detection accuracy are separately 98.9%, 94.8%, 94.4%. The accuracy of the deep learning model is 91.5. The experimental results show that the machine learning classifiers with critical features have

better accuracy than deep learning. In future work, the effect of the six features will be further verified for other classifiers.

Acknowledgements Thanks to the Fujian Provincial Key Laboratory of Big Data Mining and Applications for supporting the research.

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Ashok Kumar D, Venugopalan SR (2018) A novel algorithm for network anomaly detection using adaptive machine learning. In: Progress in advanced computing and intelligent engineering. Springer, pp 59–69
2. Awais A, Farooq M, Javed MY (2008) Attack analysis & bio-inspired security framework for ipmultimedia subsystem. In: Proceedings of the 10th annual conference companion on Genetic and evolutionary computation, pp 2093–2098
3. Boping Z (2019) Distributed svm face recognition based on hadoop. Clust Comput 22(1):827–834
4. Boser BE, Guyon IM, Vapnik V (1992) A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on computational learning theory, pp 144–152
5. Chandra K, Kapoor G, Kohli R, Gupta A (2016) Improving software quality using machine learning. In: 2016 International conference on innovation and challenges in cyber security (ICICCS-INBUSH). IEEE, pp 115–118
6. Choudhary S, Kesswani N (2020) Analysis of kdd-cup'99, nsl-kdd and unsw-nb15 datasets using deep learning in IoT. Procedia Computer Science 167:1561–1573
7. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
8. Cristianini N, Shawe-Taylor J et al (2000) An introduction to support vector machines and other kernel-based learning methods, Cambridge University Press, Cambridge
9. Dash M, Liu H (1997) Feature selection for classification. Intell Data Anal 1(1-4):131–156
10. Deebak BD, Muthaiah R, Thenmozhi K, Swaminathan PI (2013) Ip multimedia subsystem—an intrusion detection system. SmartCR 3(1):1–13
11. Deng X, Li Y, Weng J, Zhang J (2019) Feature selection for text classification: A review. Multimed Tools Appl 78(3):3797–3816
12. Drury B, Valverde-Rebaza J, Moura M-F (2017) Alneu de Andrade a Lopes: A survey of the applications of bayesian networks in agriculture. Eng Appl Artif Intel 65:29–42
13. Fratello M, Tagliaferri R (2018) Decision trees and random forests. Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics, pp 374–383
14. Freund Y, Mason L (1999) The alternating decision tree learning algorithm. In: icml, vol 99. Citeseer, pp 124–133
15. Huaiguang Wu, Li D, Cheng M (2019) Chinese text classification based on character-level cnn and svm. Int J Intell Inf Database Sys 12(3):212–228
16. Huang D-S (1996) Systematic theory of neural networks for pattern recognition. Publishing House of Electronic Industry of China, Beijing, 201
17. Huang D-S (1999) Radial basis probabilistic neural networks: Model and application. Int J Pattern Recognit Artif Intell 13(07):1083–1101
18. Huang D-S (2004) A constructive approach for finding arbitrary roots of polynomials by neural networks. IEEE Trans Neural Netw 15(2):477–491
19. Huang D-S, Ip HH-S, Law KCK, Chi Z (2005) Zeroing polynomials using modified constrained neural network approach. IEEE Trans Neural Netw 16(3):721–732
20. Kalousis A, Prados J, Hilario M (2007) Stability of feature selection algorithms: a study on high-dimensional spaces. Knowl Inform Syst 12(1):95–116
21. Kohavi R, John GH (1997) Wrappers for feature subset selection. Artif Intell 97(1-2):273–324
22. Kumar N, Kharkwal N, Kohli R, Choudhary S (2016) Ethical aspects and future of artificial intelligence. In: 2016 International conference on innovation and challenges in cyber security (ICICCS-INBUSH). IEEE, pp 111–114

23. Marcot BG, Penman TD (2019) Advances in bayesian network modelling: Integration of modelling technologies. Environ Modell Softw 111:386–393
24. Marques O, Baillargeon P (2005) A multimedia traffic classification scheme for intrusion detection systems. In: Third international conference on information technology and applications (ICITA'05), vol 2. IEEE, pp 496–501
25. Meng S, Huang W, Yin X, Khosravi MR, Li Q, Wan S, Qi L (2020) Security-aware dynamic scheduling for real-time optimization in cloud-based industrial applications. IEEE Transactions on Industrial Informatics
26. Molina LC, Belanche L, Nebot À (2002) Feature selection algorithms: A survey and experimental evaluation. In: 2002 IEEE international conference on data mining, 2002. Proceedings. IEEE, pp 306–313
27. Naomi S (1992) Altman. An introduction to kernel and nearest-neighbor nonparametric regression. Am Stat 46(3):175–185
28. Platt J (1998) Sequential minimal optimization: A fast algorithm for training support vector machines
29. Qi L, Hu C, Zhang X, Khosravi MR, Sharma S, Pang S, Wang T (2020) Privacy-aware data fusion and prediction with spatial-temporal context for smart city industrial environment. IEEE Transactions on Industrial Informatics
30. Rahnema M (2008) The core network technologies, design and dimensioning
31. Rajesh S, Paul V, Menon VG, Khosravi MR (2019) A secure and efficient lightweight symmetric encryption scheme for transfer of text files between embedded iot devices. Symmetry 11(2):293
32. Rish I et al (2001) An empirical study of the naive bayes classifier. In: IJCAI 2001 workshop on empirical methods in artificial intelligence, vol 3, pp 41–46
33. Rosenberg J, Schulzrinne H, Camarillo G, Johnston A, Peterson J, Sparks R, Handley M, Schooler E et al (2002) Sip: session initiation protocol
34. Ross Quinlan J (1987) Generating production rules from decision trees. In: ijcai, vol 87. Citeseer, pp 304–307
35. Ross Quinlan J (1996) Learning decision tree classifiers. ACM Computing Surveys (CSUR) 28(1):71–72
36. Russell SJ, Norvig P (2002) Artificial intelligence: a modern approach. 2nd edn
37. Tan P-N, Steinbach M, Kumar V (2016) Introduction to data mining. Pearson Education India, Pearson
38. Trabelsi A, Elouedi Z, Lefevre E (2019) Decision tree classifiers for evidential attribute values and class labels. Fuzzy Set Syst 366:46–62
39. Wu X, Khosravi MR, Qi L, Ji G, Dou W, Xu X (2020) Locally private frequency estimation of physical symptoms for infectious disease analysis in internet of medical things. Comput Commun 162:139–151
40. Zhang Y, Cao G, Wang B, Li X (2019) A novel ensemble method for k-nearest neighbor. Pattern Recogn 85:13–25

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.