An Effective Parallel SVM Intrusion Detection Model for Imbalanced Training Datasets

Jing Zhao[1,2], Jun Li[1,2], Chun Long[1,2], Jinxia Wei[1], Guanyao Du[1,2], Wei Wan[1,2] and Yue Wang[1,2]
[1]*Computer Network Information Center, Chinese Academy of Sciences, Beijing, China*
[2]*University of Chinese Academy of Sciences, Beijing, China*

Keywords: Intrusion Detection, Feature Reduction, Parallel SVM, Classification.

Abstract: In the field of network security, the Intrusion Detection Systems (IDSs) always require more research on detection models and algorithms to improve system performance. Meanwhile, higher quality data is critical to the accuracy of detection models. In this paper, an effective parallel SVM intrusion detection model with feature reduction for imbalanced datasets is proposed. The model includes 3 parts: 1) NKSMOTE-a Modified unbalanced data processing method. 2) feature reduction based on Correlation Analysis. 3) Parallel SVM algorithm combining clustering and classification. The NSL-KDD dataset is used to evaluate the proposed method, and the empirical results show that it achieves a better and more robust performance than existing methods in terms of the accuracy, detection rate, false alarm rate and training speed.

1 INTRODUCTION

In recent years, with the rapid development of the Internet, the scale of the network and the number of users have continued to expand, and security incidents caused by the network have emerged in the world. For example: A new ransomware virus, WannaCry, which appeared in May 2017, spread worldwide in a short period of time, causing adverse effects on the public infrastructure networks of governments, enterprises and individuals (Jesse M. E., 2017). On March 2, 2018, GitHub suffered a 1.35TB traffic attack. On April 17, 2019, a research report from the Cisco Talos research team stated that a state-sponsored hacker successfully hijacked the Domain Name System (DNS) records and stole certificates from nearly 40 public and private entities in 13 countries, which caused serious consequences. According to statistics from relevant departments (Source: Ponemon Institute and Juniper Networks, October 2017), 12,172 alerts are generated globally every week, and 518 alerts are studied every week. It takes 352.3 hours to track false alarms every week. Tracking false alarms costs $1,145,000 per year, and cyber breaches cost $7,000,000 per year.

With the continuous update and mutation of attack methods in the network, the traditional intrusion detection methods based on feature matching have been unable to effectively solve security problems, especially when the characteristics of attackers are becoming more and more complex, the traditional intrusion detection methods show great disadvantages: 1) low detection rate and high false alarm rate, 2) poor adaptability, 3) low detection efficiency. Therefore, intrusion detection methods based on machine learning are emerging.

Machine learning technology has been widely used in intrusion detection system for a long time. Many efficient Machine learning models are used to classify normal and abnormal traffic in intrusion detection, such as Support Vector Machine (SVM) (B M A., 2016), Decision Tree (L M., 2018), K-means (W L AL., 2017) and Artificial Neural Network (ANN) (S. S. Roy, 2017).

In a large number of researches using machine learning to solve intrusion detection problems, the imbalance between training data and test data often leads to low efficiency of algorithms. Many researches in the field of intrusion detection focused on unbalanced data sets. Hamid et al. proposed a hybrid method based on coupling discrete wavelet transforms and ANN （Artificial Neural Network） for intrusion detection (Y. H., 2018). They eliminated the imbalance of the instances by SMOTE (Synthetic Minority Oversampling Technique) based oversampling of less frequent class and random under-sampling of the dominant class. S. M. H. B et

Zhao, J., Li, J., Long, C., Wei, J., Du, G., Wan, W. and Wang, Y.
An Effective Parallel SVM Intrusion Detection Model for Imbalanced Training Datasets.
DOI: 10.5220/0009390302250232
In *Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS 2020) - Volume 2*, pages 225-232
ISBN: 978-989-758-423-7

al. introduced penalized multiple criteria linear programming (MCLP) to deal with unbalanced datasets (S. M. H. B, 2016). Aggarwal proposed a new metric NPR (NP Ratio) used for ranking the classifiers for IDS to evaluate the detection rate and false alarm rate caused by unbalanced data (P A. 2016). Wang Li proposed an improved NKSMOTE algorithm to overcome the shortage of SMOTE. A nonlinear mapping function was used to map samples to a high-dimensional kernel space (Wang Li, 2018). Such solutions to unbalanced data sets can improve the classification accuracy to some extent, but also significantly increase the complexity of model training.

No matter which model is used, the steps of machine learning algorithms for intrusion detection classification can be divided into two major parts: the data preprocessing and the model training. If the metadata and its characteristics can be reduced and selected in the data preprocessing stage, the model framework will be more concise, and the computational efficiency will be improved, and so that the generalization ability of the model will be enhanced. Yin et al. proposed a novel classification algorithm using Ensemble Feature Selections (EFS) for the imbalanced-class dataset. This algorithm introduced a penalty-reward mechanism for minority classes when feature subset objective functions were designed (H Y., 2016). In (A A A., 2016), the integration of linear discriminant analysis (LDA) and principal component analysis (PCA) feature extraction algorithm was used to alleviate the imbalance of data.

To solve the above problems, this paper first creatively propose a simplified NKSMOTE (a Modified unbalanced data processing method) to make samples of different categories nearly balanced in quantity, and then use a feature reduction algorithm to reduce feature dimensions to improve the accuracy and efficiency of the intrusion detection. Finally, a parallel SVM model was designed to realize multi-classification intrusion detection for unbalanced datasets.

2 PRELIMINARY

2.1 SMOTE Algorithm

The SMOTE algorithm is a classic oversampling algorithm (Chawla N V, 2011). This algorithm is different from the traditional oversampling method. It is not a simple copy of the sample. The main idea is to insert new artificial samples into the samples close to the minority class samples. This can increase the number of the minority class samples, and can effectively solve the problem of classification overfitting. The specific method is as follows:

Suppose S is a set of samples and X is a set of the minority class samples, where $X \subset S$, $x_i \subset X$. Find the k nearest neighbors of the same class for each x_i through the oversampling rate N and interpolate to obtain a new synthesized sample.

$$x_{new} = x_i + rand(0,1) * (y_j - x_j) \tag{1}$$

$j = 1, 2, ...n$, x_{new} represents the sample obtained after oversampling, y_j represents the k nearest neighbors of the same type in x_i, and $rand(0,1)$ represents a random number in the region $(0,1)$.

2.2 Feature Sorting based on Correlation Analysis

Definition 1. For a discrete feature vector, its probability distribution can be expressed as $\{p(x_1'), p(x_2'), \cdots, p(x_n')\}$, then entropy of feature is as follows:

$$H(X) = -\sum_{i=1}^{n} p(x_i') \log_2 p(x_i') \tag{2}$$

If all the values of X are the same, then the entropy of X is 0. Thus, the feature X is useless for data classification.

Definition 2. For two discrete features $X \in \{x_1', x_2', \cdots, x_n'\}^T$ and $Y \in \{y_1, y_2, \cdots, y_m\}$, their joint probability density is $p(x_i', y_j), 1 \le i \le n, 1 \le j \le m$, and conditional density is $p(x_i' | y_j)$, then entropy of X under the condition Y can be expressed as

$$H(X|Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i', y_j) \log_2 \frac{p(y_j)}{p(x_i', y_j)} \tag{3}$$

The mutual information is generated and derived from entropy. For two features X and Y in one dataset, the mutual information between them is as follows:

$$\begin{aligned} & I(X;Y) \\ & = H(X) - H(X|Y) \\ & = \sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i', y_j) \log_2 \frac{p(x_i', y_j)}{p(x_i') p(y_j)} \end{aligned} \tag{4}$$

The mutual information has the following characteristics:

Symmetry: $I(X;Y) = I(Y;X)$

Monotonic: if $A \subseteq B \subseteq C$, then $I(A;C) = I(B;C)$

The mutual information reflects the amount of information shared between two random variables. The greater value of the mutual information, the greater correlation between the two variables. If the mutual information between two variables is 0, the two variables are completely uncorrelated and statistically independent in probability.

3 RESEARCH METHODOLOGY

This section introduces the research methodology for intrusion detection model , which includes three parts: NKSMOTE-a modified unbalanced data processing method, feature reduction based on correlation analysis and intrusion detection algorithm based on Online-SVM.

3.1 S-NK-SMOTE: A Modified Unbalanced Data Processing Method

The purpose of S-NK-SMOTE algorithm is to balance the number of different categories samples. Our method is a simplification of the NK-SMOTE (Wang Li, 2018). the processing steps are as follows:

Step1 : For all minority class samples x , search its nearest k samples. Among the k samples, if the number of minority class samples is greater than the number of majority class samples, x is a safety sample; if the number of minority class samples is less than the number of majority class samples, and if exist minority class samples, x is a dangerous sample; if all samples are majority class samples, then x is a noise sample.

Step2 : If x is a noise sample, oversampling will introduce the risk of noise into the dataset, while the noise samples have certain positive impact on the classification. In order to reduce the risk of noise, select a sample x' randomly from the minority class, and generate a new sample close to the minority class:

$$X_{new} = x + rand(0.5,1) \cdot (x' - x) \qquad (5)$$

Step3 : If x is not a noise sample, then select 1 sample x' randomly from its k nearest samples. If x' belongs to the samples of the majority class, a new sample close to x is generated as follows:

$$X_{new} = x + rand(0,0.5) \cdot (x' - x) \qquad (6)$$

if x' belongs to the samples of the minority class, a new sample close to x is generated according to the following formula:

$$X_{new} = x + rand(0,1) \cdot (x' - x) \qquad (7)$$

3.2 Feature Reduction based on Correlation Analysis

After balancing the dataset, high dimensional samples can easily increase the classifier training time and computational complexity, meanwhile, overfitting will inevitably occur. Therefore, representing a subset of features with effective lower dimension is significant to improving classification accuracy and efficiency.

This paper will calculate the correlation between each feature and classification label and select the best feature. Detailed steps are as follows:

Step1 : According to formulas (2) (3) (4), calculate the correlation between each feature vector $X \in \{x_1', x_2', \cdots, x_n'\}^T$ (n is the number of samples) and its corresponding category label $Y \in \{0, 1\}$, and perform standardized compression, the correlation is expressed as: $C(X,Y) = \frac{I(X;Y)}{H(X)+H(Y)}$, where $0 \le C(X,Y) \le 1$,if $C(X,Y)$=1 , indicating features are fully related to category labels, if $C(X,Y)$=1 , indicating features are fully unrelated to category labels.

Step2 : Sort the features from large to small according to the correlation. The higher the correlation, the more information the feature contains, the more useful it is, and it can more accurately reflect the specific situation of the sample.

Step3 : SVM algorithm is used for feature selection. Considering the accuracy of intrusion detection is closely related to feature selection, ACC is used as the criterion for feature selection. The optimal solution lies in fewer features and higher accuracy, specific algorithm is as follows:

Suppose we have a dataset with m features, the features have been sorted from large to small according to correlation, denoted as R_0 , the accuracy ACC_0 of R_0 is calculated by SVM. Remove the last feature of R_0 to generate a new dataset, denoted as R_1 , the accuracy ACC_1 of R_1 is calculated by SVM, if $ACC_1 > ACC_0$ then keep R_1 , otherwise discard it.

Remove the last feature of R_1 to generate a new dataset, denoted as R_2, and so on remove the last $t-th$ feature of R_t, if $ACC_t > ACC_{t-1}$ then keep R_t otherwise discard it. Finally, the dataset corresponding to the optimal ACC and the most representative features are selected.

3.3 Intrusion Detection Algorithm based on Online-SVM

In this section, we will design an Online-SVM based detection algorithm that can achieve multiple classification of abnormal samples. The algorithm includes a training phase and a detection phase.

3.3.1 Training

When training the algorithm, combine each class of abnormal data with normal data to generate a synthesized dataset, and then train an Online-SVM classifier on the synthesized dataset. Assume that there exist m different (sub)classes of abnormal data, denoted as $\{Att_1, Att_2, \cdots, Att_m\}$, we can get m new synthesized training set and m Online-SVM classifier. Finally, combine all abnormal data with normal data to generate the $(m+1)th$ synthesized training set, and train the $(m+1)th$ Online-SVM classifier. Detailed steps during the training phase are as follows:

Step1 : Use the methods provided in section 3.1 and section 3.2 to process m abnormal sample sets $\{Att_1\_T, Att_2\_T, \cdots, Att_m\_T\}$ and 1 normal sample set $\{Normal\}$.

Step2 : Generate $m+1$ synthesized training sets, $\{Att_1\_T, Normal\}$, $\{Att_2\_T, Normal\}$, $\cdots$, $\{Att_m\_T, Normal\}$ and $\{Att_1\_T, Att_2\_T, \cdots, Att_m\_T, Normal\}$.

Step3 : Train $m+1$ Online-SVM classifiers, Clf_Att_1, Clf_Att_2, $\cdots$, Clf_Att_m and Clf_Normal respectively.

3.3.2 Testing

The intrusion detection algorithm consists of $m+1$ Online-SVM classifiers, and the training sets each classifier used are different. Therefore, it can be customized and expanded flexibly.

In order to improve the classification efficiency, K-means is used to determine which classifier the testing data input into. Specifically, we use K-means to calculate the cluster center of all categories of samples, and then calculate the average distance between the testing data and each cluster center. At last we choose the Silhouette Coefficient to evaluate the result. The Silhouette Coefficient mainly measures the similarity of a sample and other samples of its own cluster, and the difference with the sample and other samples of other clusters. The higher the Silhouette Coefficient, the better the determine result. Detailed methods are as follows:

Step1 : Calculate the average distance between sample x and other samples in the same cluster, called the Dissimilarity of the cluster, expressed as $a(x)$.

Step2: Calculate the average distance between sample x and the other samples in other H clusters, expressed as $\{b_{x1}, b_{x2}, \cdots, b_{xH}\}$, then the Inter-cluster Dissimilarity can be expressed as $b(x) = \min\{b_{x1}, b_{x2}, \cdots, b_{xH}\}$.

Step3: According to $a(x)$ and $b(x)$, define the Silhouette Coefficient of sample x:

$$s(x) = \frac{b(x)-a(x)}{\max\{a(x), b(x)\}} = \begin{cases} 1-\dfrac{a(x)}{b(x)}, & a(x) < b(x) \\ 0 & , a(x) = b(x) \\ \dfrac{b(x)}{a(x)} - 1, & a(x) > b(x) \end{cases} \quad (8)$$

Step4: The Silhouette Coefficient of the current clustering result is the average of all samples $s(x)$, and the value range is $[-1,1]$, the closer to $+1$, the better the determine result.

Specific detection process is as follows:

Let $K(i)$ be the number of cluster centers in class i, and the $k-th$ cluster center in class i be expressed as C_i^k, the specific steps are as follows:

Step1 : Calculate the average distance $dist(i)$ between the sample x to each class:

$dist(i) = \sum_1^{K(i)} \sqrt{(x - C_i^k)^2} \,/\, K(i), i \in \{1, 2, \ldots, m+1\}$.

Step2 : Input x to $Clf-ind$, $ind = \arg\min_i dist(i)$.

Step3 : If $ind \in \{Att_1, Att_2, \cdots, Att_m\}$, the output is $cls \in \{Att_{ind}, Abnormal\}$. Att_{ind} represents abnormal type.

Step4 : If $ind = Normal$, the output is $cls \in \{Att_{unknown}, Normal\}$. It represents two result: x is a normal sample, or x can be considered as an unknown attack.

4 EXPERIMENTS

4.1 Experimental Data and Environment

In this section, the experimental results on several simulation and real methods are discussed. The dataset used in the paper is based on the NSLKDD, which is a modified version of the KDD Cup 99 (Tavallaee, 2009) dataset. The NSL-KDD testing dataset contain 22544 connection patterns. Each testing dataset have more number of attacks than that of in training dataset.

The empirical experiments in our work were all implemented on a computer with an Intel Core i7-7700 CPU @ 3.60GHz with 16.0 GB RAM running Windows 10. The proposed intrusion detection algorithm was implemented by using Python.

4.2 Experimental Results and Analysis

The parallel SVM algorithm aims to improve the accuracy for imbalanced datasets. Therefore, to evaluate the performance of our proposed models and make expedient comparisons with other detection methods, we chose accuracy (ACC), detection rate (DR) and false alarm rate (FAR) as the evaluation measures. These three measurements are calculated from the following formula.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$DR = \frac{TP}{TP + FP} \tag{10}$$

$$FAR = \frac{FP}{FP + TN} \tag{11}$$

To verify the effectiveness of our proposed intrusion detection models, we first discuss three situations in this section: 1) Using Online-SVM algorithm only. 2) Using NKSMOTE before Online-SVM. 3) Using NKSMOTE and feature reduction before Online-SVM. In order to evaluate the performance of our model, all the experiments were executed under the 5-class classification, and 10-fold cross-validation was utilized to reduce the risk of overfitting. The experimental results are shown in Table 1-3.

Table 1: Performances of 5-class classification using Online-SVM only.

Attack Types	Accuracy (%)	DR (%)	FAR (%)
Normal	98.89	98.51	0.9
DOS	97.5	97.62	0.98
Probing	98.21	98.1	0.9
R2L	98.16	97.78	0.88
U2R	97.83	97.4	0.7
Total	98.12	97.87	0.9

Table 2: Performances of 5-class classification using NKSMOTE before Online-SVM.

Attack Types	Accuracy (%)	DR (%)	FAR (%)
Normal	99.4	99.03	0.65
DOS	98.62	98.49	0.55
Probing	98.98	98.85	0.66
R2L	98.96	98.77	0.512
U2R	98.72	98.8	0.43
Total	98.89	98.8	0.527

Table 3: Performances of 5-class classification using NKSMOTE and feature reduction before Online-SVM.

Attack Types	Accuracy (%)	DR (%)	FAR (%)
Normal	99.55	99.43	0.5
DOS	98.32	98.9	0.16
Probing	99.56	99.51	0.23
R2L	99.33	99.43	0.55
U2R	99.18	99.46	0.3
Total	99.53	99.51	0.348

Table 4: Performance comparison of different intrusion detection methods with NSL-KDD.

Literature	Method	Accuracy (%)	DR (%)	FAR (%)
Our method	Parallel SVM	99.53	99.51	0.348
A M V. B et al.	MCC-MCLP	99.13	99.09	0.55
Y. et al.	MK-SVM	98.63	98.97	0.66
Bo Hu et al.	DDF	92.70	91.70	0.50

Table 5: Run time of different intrusion detection methods.

Literature	Method	Training time	Testing time
Our method	Parallel SVM	103s	10ms
Bo Hu et al.	DDF	256s	10ms
Chen et al.	XGBoost	198.2s	5.2ms
A. et al	Deep Belief Network	27330s	11ms

In Table 1, only parallel SVM algorithm is applied in NSL-KDD. The values of ACC, DR and FAR under different feature conditions are illustrated. From the result in Table 1, we can see that the performance of each classification is stable and relatively excellent. The result means parallel SVM algorithm is suitable for multi-classification.

In Table 2, the NKSMOTE is applied to process unbalanced data, forming new dataset that the quantity of different categories tends to balance. Then parallel SVM algorithm is applied in processed NSL-KDD. From the result in Table 1 and Table 2, we can get a conclusion that using NKSMOTE algorithm is very effective. The values of ACC and DR increase much and the value of FAR decreases obviously.

In Table 3, the NKSMOTE is applied to process unbalanced data, and the proposed methods mentioned in 3.2 is used to perform feature reduction. The results show that when the feature is reduced to 25 dimensions the value of ACC is optimal. Here the 25-dimensional subset consists of feature as <6, 3, 35, 27,4, 26, 23, 33, 38, 5, 34, 24, 39, 29, 36, 12, 37, 24, 32, 2, 40, 31, 1, 27, 28>. From the result in Table 3, we can get a conclusion that Using NKSMOTE and feature reduction simultaneously, the results are optimal. However, Comparing the results of two tables above, it is found that the NKSMOTE algorithm is more effective.

To verify the effectiveness of our proposed intrusion detection models, we compare the performance of MCC-MCLP (A M V. B., 2017), MK-SVM (Y., 2012) and Dynamic Deep Forest (DFF) (Bo Hu, 2019) on the NSL-KDD dataset. Table 4 and Figure 1 show the performance of different intrusion. The results illustrate that our method provides the highest accuracy rate of 99.53%, the highest detection rate of 99.51%, and the lowest false alarm rate of 0.348%. The comparison results show that the proposed method is superior to other intrusion detection scheme in detection performance.

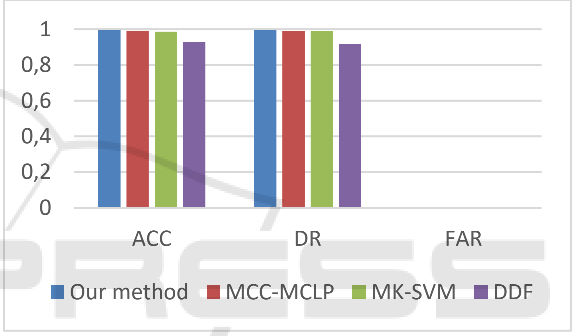

Figure 1: Performance comparison of different intrusion detection models with NSL-KDD.

Table 5 and Figure 2 shows the results obtained by comparing the run time of different intrusion detection methods. The performance of our method is not very well in testing time.

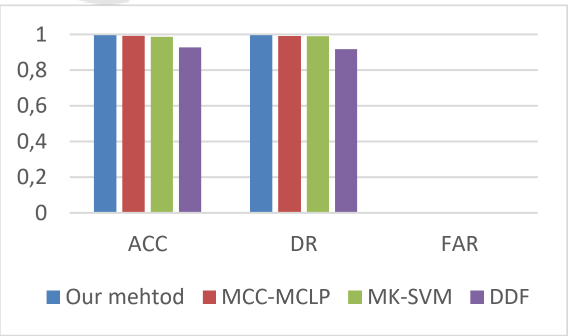

Figure 2: Run time of different intrusion detection methods.

5 CONCLUSIONS

In order to improve the accuracy of the intrusion detection for imbalanced training datasets, this paper proposes an effective model, which include 3 parts: 1) NKSMOTE-a modified unbalanced data processing method. 2) feature reduction based on correlation analysis, which reduces the original 41-dimensional feature to a 25-dimensional feature. 3) parallel SVM algorithm combining clustering and classification.

The experimental results illustrate that our proposed detection method can obtain an outstanding performance with a high ACC, a high DR, a low FAR and a rapid training speed. However, the results show that our method has no advantage in testing time comparing with other newly related works. Therefore, we will consider to improve the performance of the proposed method in running time.

ACKNOWLEDGEMENTS

The authors would like to thank the editorial board and reviewers. This work was supported by the Construction Project of Network Security System (XXH13507); the Strategic Priority Research Program of Chinese Academy of Sciences (XDC02030600).

REFERENCES

Jesse M. E., 2017. WannaCry, Cybersecurity and Health Information Technology: A Time to Act. *Journal of Medical Systems, 41(7).*

B M A., R R., M C., et al, 2016. A hybrid method consisting of GA and SVM for intrusion detection system[J]. *Neural Computing and Applications, 27(6):1669-1676.*

L M., P S., N V D., 2018. A Data Classification Model: For Effective Classification of Intrusion in an Intrusion Detection System Based on Decision Tree Learning Algorithm[J]. *Springer, Singapore. 61-66.*

W L AL., Z A O., M Z A N., 2017. Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system[J]. *Expert Systems with Applications*, *67:296-303.*

S. S. Roy, A. M., R. G., M. S., Obaidat and P. V. K., 2017. A Deep Learning Based Artificial Neural Network Approach for Intrusion Detection. International Conference on Mathematics and Computing[C]. *International Conference on Mathematics and Computing.* Springer.

Y. H., F. A. Shah and M. S., 2018. Wavelet neural network model for network intrusion detection system[J]. *International Journal of Information Technology.*

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

P A., S K P., 2016. A Metric for Ranking the Classifiers for Evaluation of Intrusion Detection System[J]. *Proceedings of the Second International Conference on Computer and Communication Technologies. Advances in Intelligent Systems and Computing, vol 380.* Springer, New Delhi.

Wang Li, 2018. NKSMOTE Algorithm Based Classification Method for Imbalanced Dataset. *Computer Science. Vol. 45 No.9.*

Han, H., Wang, W., Mao, B., 2005. Borderline−SMOTE: A New Over−Sampling Method in Imbalanced Data Sets Learning. *International Conference on Intelligent Computing:878-887.*

H Y., K G., Z W., 2016. A Classification Algorithm Based on Ensemble Feature Selections for Imbalanced-Class Dataset[C]. *IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS).*

A A A., M B I R., 2016. Ensemble of binary SVM classifiers based on PCA and LDA feature extraction for intrusion detection[C]. *IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC).*

Chawla N V, Bowyer K W, Hall L O, et al., 2011 SMOTE: Synthetic Minority Over-sampling Technique[J]. *Journal of Artificial Intelligence Research, 16(1):321-357.*

Wang G P, Yang J X, Li R, 2017. Imbalanced SVM-Based Anomaly Detection Algorithm for Imbalanced Training Datasets[J]. *ETRI Journal, 39(5):621-631.*

P. Lakshmi and D. Geetha, 2016. Intrusion Detection System using Modified Support Vector Machine[J]. *International Research Journal of Engineering Science and Technology, Vol. 1, No. 1.*

Jinxia Wei, Chun Long, Wei Wan, Yurou Zhang, Jing Zhao and Guanyao Du, 2019. An Effective RF-based Intrusion Detection Algorithm with Feature Reduction and Transformation[C]. *ICEIS 2019.*

Aman Mudgal1, Rajiv Munjal2., 2012. Fuzzy K-Means Based Intrusion Detection System Using Support Vector Machine. *International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Impact Factor (2012): 3.358.*

Huiwen, Wang, Jie, et al, 2017. An effective intrusion detection framework based on SVM with feature augmentation[J]. *Knowledge-Based Systems.*

A D L., 2016. A hybrid approach to reducing the false positive rate in unsupervised machine learning intrusion detection[C]. *SoutheastCon 2016.* IEEE.

A B., Guven E, 2015. A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection[J]. *IEEE Communications Surveys & Tutorials, 18(2):1-1.*

B I., Yadav A, Soni A K, 2017. Decision Tree Based Intrusion Detection System for NSL-KDD Dataset[C]. *International Conference on Information and Communication Technology for Intelligent Systems (ICTIS 2017). Springer, Cham.*

Tavallaee M., Bagheri E., Lu W., Ghorbani A. A., 2009. A detailed analysis of the KDD CUP 99 data set[C], Tavallaee M , Bagheri E , Lu W , et al. A detailed analysis of the KDD CUP 99 data set[C]. *IEEE International Conference on Computational Intelligence for Security & Defense Applications.*

A M V. B., A M.B., 2017. A multi-class classification MCLP model with particle swarm optimization for network intrusion detection [J]. *Sadhana: Academy Proceedings in Engineering Science.*

Y., L., Weidong, L., Guoqiang, W., 2012. An Intrusion Detection Approach Using SVM and Multiple Kernel Method[J]. *IJACT: International Journal of Advancements in Computing Technology 4(1), 463–469 (2012).*

Bo Hu, Jinxi Wang, 2019. Dynamic Deep Forest: An Ensemble Classification Method for Network Intrusion Detection. *Electronics 2019, 8(9), 968.*

Chen, T.Q.; Guestrin, C, 2016. XGBoost: A Scalable Tree Boosting System[C]. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM.*

A., K.; Purdy, C, 2017. Toward an online anomaly intrusion detection system based on deep learning[C]. *Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2017; pp. 195–200.*