DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

Intrusion Detection System Based on Ada boosting and Bagging Algorithm

Ali K. Hilool[1], Soukaena H. Hashem[2], Shatha Habeeb [3]
[1,2,3]*Department of Computer Science, University of Technology, Baghdad, Iraq*
[1]*Ali1995a1995a@gmail.com*, [2]*110015@uotechnology.edu.iq* ،[3]*110056@uotechnology.edu.iq*

Abstract— *Computer worms execute damaging functions in the network systems, compromising system security. Although researchers use a variety of methods to detect worms and prevent their spread. Detecting worms remains a challenge for the following reasons: First, a huge volume of irrelevant data affects classification accuracy. Second, frequently used individual classifiers in systems are poor at detecting all types of worms, Third, many systems are built on out-of-date information, rendering them useless for new worm species. As a result, providing a network intrusion detection system is vital for ensuring security and reducing the harm caused by worms on networks to information systems. The goal of the study is to discover computer worms in the computer networks and protect the systems from their damages. The proposed method uses the UNSW NB15 dataset to train and test the ensemble Ada boosting and Bagging algorithms with the Support Vector Nachine (SVM) as a contribution rather than a decision tree. Due to Correlation Feature Selection (CFS) identifying relationships between features and classes, and Chi-square (Chi2) determining whether features and classes are independent or not, we combined these two algorithms as a contribution in a method called CFS&Chi2fs to select the relevant features and reduce the time. The system achieved accuracy reaching 0.998 with Bagging(SVM), and 0.989 with Ada boost(SVM).*

Index Terms— *Intrusion Detection System, Computer Worms, Ada boosting, Bagging.*

I. INTRODUCTION

The internet has turned our world on its head. It has revolutionized communications to the point where it is now our preferred mode of communication daily. Because of the increased use of the internet, cyber attacks have emerged, which are unwanted attempts to steal, expose, alter, disable, or destroy information by gaining unauthorized access to computer systems. Computer worms, for example, are a common type of cyber attack that can shut down networks or steal data. Computer worms are small, self-contained programs that do not require any external assistance [1]. They're made to conduct out malicious operations, steal data from users while they're perusing the web, or harm them or their callers. They spread swiftly and are difficult to eradicate due to their exceptional capacity to avoid detection, multiply, and colorize, because it infects machines connected to the network automatically and without human intervention, the worm spreads more extensively and quicker than viruses [2]. Several methods have been used to detect computer worms and estimate their impact, including the use of machine learning techniques, encryption, a firewall, and a variety of other methods [3]. Intrusion Detection System (IDS) is one of the most reliable systems for detecting penetrations and attacks, it is software that identifies all activity, whether it is normal or malignant [4]. IDS generates a lot of false alarms, and this problem has prompted a lot of academics to try to figure out how to separate False Positive

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

(FP) and False-Negative (FN) alarms and reduce false alarms (FN). Data mining-based IDS can be used to improve IDS in real time, remove routine activity from alarm data to focus on real assaults, and detect aberrant activity that reveals a true attack. It's a computational framework that combines database systems artificial intelligence, and machine learning to detect patterns in multiple datasets. Data mining applications can employ a variety of characteristics to examine diverse data sets [5], [6]. Due to the consequences of rising security threats nowadays, Network Intrusion Detection Systems (NIDS) have become the most crucial component of modern network infrastructure. Although the Intrusion Detection System (IDS) generates a lot of alarms, it uses algorithmic processes to limit false positives [7], [8], and [9]. Ensemble learning is a machine learning technique that entails teaching a bunch of weak learners (models) to solve a problem and then combining their findings to get superior results. The core premise is that we can get more accurate and/or robust models by merging weak models properly. The three types of ensemble approaches are as follows. Bagging is a way of combining homogeneous weak learners, training and testing them simultaneously, and then combining them using voting, average, and other approaches. Boosting, which brings together a group of similar weak learners and trains and tests them in a systematic manner (each iteration depends on previous ones). Stacking is an ensemble strategy in which a new model learns the most efficient way to combine the predictions of multiple current models [10]. The contribution of this paper is to combine the results of correlation and chi-square feature selection and use a support vector machine classifier instead of the decision tree in the Ada boost and bagging algorithms, where the decision tree is the default base classifier. The rest of the paper is organized as follows: The related work of worm identification is discussed in Section II. The theoretical underpinning of the model is provided in Section III. The worm detection system architecture, which is based on ensemble Bagging and Ada boosting, is introduced in Section IV. In Section V, we detail our extensive tests for evaluating the proposed worm detection system. Section VI wraps up by elaborating on the conclusion.

II. RELATED WORK

In [11] Pelin Yildirim Taser suggested the bagging and boosting strategy based on six Decision Tree-Based (DTB) which are C4.5, Random Tree, REPTree, Decision Stump, Hoeffding Tree, and Naive Bayes Tree was employed to predict diabetes. A comparison of individual implementation, boosting, and bagging of DTB classifiers is done by the author in terms of accuracy rates. Experimental results demonstrate that Ada boost with Naive Bayes Tree (NBTree) has the best accuracy score of 98.65 percent. In [12] Hoang Ngoc Thanh and Tran Van Lang suggested a fuzzers detection system using the UNSW-NB15 dataset, to analyze and evaluate the experimental outcomes. Single classifiers such (J48 (DT), logistic (LR), Lib SVM (SVM), Naive Bayes (NB), Random Tree (RT), and ibk (KNN)) were used in bagging, Ada boost, stacking, random forest, and decorate ensemble techniques. The experimental results show that the Ada boost technique with component classifiers using decision tree has the best classification quality with an F-Measure of 96.76 percent, compared to 94.16 percent for single classifiers and 96.36 percent for Random Forest. In [13] Sarah Mohammad suggested a NIDS consist of three levels, in the first level, the Naive Bayes algorithm is used to distinguish abnormal behavior from normal behavior; in the second level, the Multinomial logistic regression algorithm is used to classify abnormal activity into the main four attack types; and in the third level, the ID3 Decision Tree algorithm is used to classify four attack types into (23). The (KDDCUP99) dataset will

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

be used to evaluate the proposed system, as well as entropy, is used as feature selection. The experimental results show that the accuracy reached 98 percent in the first level when 41 features were selected, 97 percent in the second level when Multinomial logistic regression was used and 41 features were selected, and 97 percent in the third level when ID3 Decision Tree was used and 41 features were selected. In [14] Dishan Jing and Hai-Bao Chen using the UNSW-NB-15 dataset and a Support Vector Machine (SVM) classifier, they proposed an intrusion detection system (IDS). There was no feature selection method used by the authors. They used nonlinear scaling instead of the min-max normalization algorithm in the preprocessing stage, claiming that it produces better results with the UNSW-NB-15 dataset. According to the findings, the accuracy in detecting attacks was (85%), while the false alarm rate was (15.26). In [15] Shigeyuki H., et, looked at defaulted loans in a Taiwanese database and compares the prediction accuracy and classification ability of three ensemble learning methods, bagging, random forest, and boosting, with that of various neural-network methods, each of which has a different activation function. The results show that boosting outperforms other machine-learning methods, such as neural networks, in terms of classification ability. The choice of activation function, the number of middle layers, and the inclusion of dropouts all affect the performance of neural network models. In [16] Ankita and Nabizath Saleenaa proposed a system for Twitter sentiment analysis based on an ensemble classification. The system combines several base classifiers such as NB, SVM, RF, and LR to produce a robust classifier and improve the performance and accuracy of base learning techniques. The algorithm has been trained and tested using a dataset gathered from Twitter called Health Care Reform (HCR). The suggested ensemble classifier outperformed stand-alone classifiers and the widely used majority voting ensemble classifier, according to the findings.

III. THEORETICAL BACKGROUND

In this section, we'll go over the theoretical aspects of the algorithms used in the articles, as follows:

1- Feature Selection: one of the most significant preprocessing processes in data mining approaches is feature selection, which is used to remove superfluous and redundant features from the dataset, enhance the model's performance by utilizing the proper features, and reduce the amount of time it takes to analyze the data. In this paper, we utilized correlation features selection and chi-square features selection.

2- Correlation Feature Selection (CFS): Correlation-based feature selection ranks characteristics by using a heuristic assessment function based on correlations. The function compared attribute vector subsets that are linked `ss features, on the other hand, should be looked at because they are typically connected with one or more of the other traits and stay only one of them. The criterion for assessing a subset of N features is as follows [17].

$$M_s = \frac{N\overline{t_{cf}}}{\sqrt{N+N(N-1)\overline{t_{ff}}}} \qquad (1)$$

MS represents the evaluation of a subset of S that has N characteristics. ($\overline{t_{cf}}$) is the average correlation value between attributes and class labels.($\overline{t_{ff}}$) is the average of the correlation between two characteristics [17]

a- Chi-Square Feature Selection

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

The Chi2 test evaluates the relationship between a class and a feature, allowing it to discover features that are more relevant for a given dataset. As a result, features that aren't relevant for categorization can be removed from the feature space [18]. From the data of two features, we will get the observed count X and projected count Y. The Chi-Square test is used to assess how far expected count Y and observed count X differ.

$$X_c^2 = \frac{(A_i - E_i)^2}{E_i} \qquad (2)$$

Where C is the degree of freedom, X denotes the observed value(s), and Y denotes the expected value (s).

We compare the value of X c2 to the value of the chi2 table value where alpha =0.05 and remove the feature if it is less than the chi2 table value (independent); else, the feature is accepted.

1- Bagging classifiers: are ensemble meta-estimators that fit base classifiers to random subsets of the original dataset and then aggregate their individual predictions (either by average or voting) to create a final prediction. By incorporating randomness into the building method of a black-box estimator (e.g., SVM), such a meta-estimator may generally be used to minimize the variance of a black-box estimator (e.g., a decision tree). Each base classifier is trained in parallel using a training set that is produced by randomly selecting N instances (or data) from the original training dataset, with replacement – where N is the size of the original training dataset. Each base classifier's training set is distinct from the others. Many of the original data points will likely be repeated in the final training set, while others will likely be omitted [19], [20].

2- Ada boost: This is a machine learning algorithm that begins by giving equal weight to all instances in the training dataset; the learning algorithm is then used to create a classifier for this data by making the number of stumps (nodes with two leaves) equal to the number of features; and finally, only one stump is chosen after calculating Gini and Entropy for all trees, that has the lowest value (Gini or Entropy). The total error (TE) and stump performance were then determined. We must update the weights for all instances of the dataset based on the stump's performance, increasing the weight for misclassified records and decreasing the weight for properly classified records. Then, using normalized weights, a new dataset is produced. The algorithm will construct a new stump based on this new dataset and continue the process until it goes through all trees sequentially and determines that there is less inaccuracy than the normalized weight that we had in the beginning stage [21], [22]

3- Instead of stumps, I recommend utilizing an SVM classifier as an update to the Ada boost method, using the same stages after creating stumps.

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

IV. RESEARCH METHODS

The effective data mining model for worm detection was proposed, which uses both anomaly and misuse detection techniques, where each case in a dataset is labeled as "attack" or "normal" (worms are one type of attack), and a learning algorithm is trained over the class data to improve worm detection in networks. *Fig. 1* depicts the structure of the suggested worm detection model. The process is divided into four stages:

1- Preprocessing of the data set: preprocessing procedures were initially added to the basic datasets to prepare the data for the classification algorithm.
2- Dimensionality reduction: A feature selection approach based on chi-square and correlation features selection is used to identify the most essential features and reduce the dimensionality of the dataset.
3- Classifier training: the Ada boost and bagging algorithms were used to build classifiers to increase the accuracy of worm identification.
4- A classification to predict the outcome of our model (testing) was utilized.

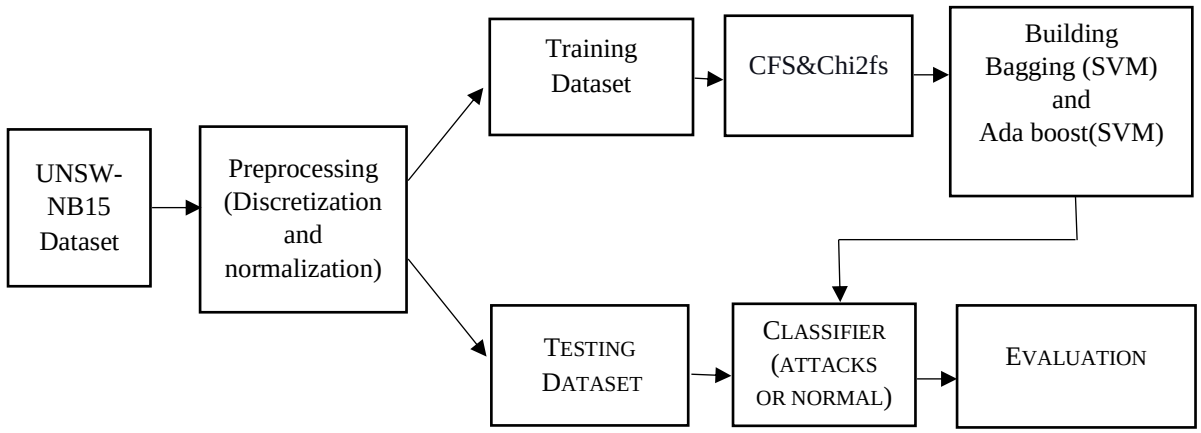

FIG. 1. FLOWCHART OF THE PROPOSED SYSTEM.

1- Preprocessing

Because the UNSW-NB15 dataset comprises both continuous and discrete characteristics, it is required to transform the continuous attributes to discrete to assure the system's efficiency and to address the issue of new values emerging in the test dataset that are not present in the training dataset. Following discretization, we utilized the Min-Max normalization method to increase the model's efficiency and efficacy by putting attribute values between 0 and 1. After discretization and normalization, we will utilize correlation feature selection and chi-square feature selection to remove unused and redundant features from the dataset See Algorithm 1.

Algorithm (1) Preprocessing
input : observations of the unsw-nb15 Datase Output: dataset ranging from zero to one, relevant features
Start Step 1: min-max normalization For each feature in the Dataset Find the maximum amount in feature

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

Find the minimum amount in feature
For each (N) rate in Feature extract

$$Values\_of(N) = \frac{Value(N) - Min}{Max - Min}$$

End_For
End_For
Step 2: correlation CFS
For each class and features
Set the correlation of class label with all features
Pick the features which have a strong relationship with class label.
Elimiate the remainder features
End_For
For each feature in the you've chosen subset
Extract the correlation of feature with all features
If two or more features are correlated
Pick one of them`and Remove the remainder
End_For

Step 3: Chi_square feature selection
For each feature in the dataset
Extract $X_c^2$with class. See the equation 1
Set alpha=0.5
From special chi2 table find X_c^2' where alpha=0.05 and compare the result to X_c^2
If $X_c^2 < X_c^{2'}$ the feature is dropeded (independent)
Else it is not drop (depended on the class)
End_For
End

2- Training and Testing

The dataset will be divided into two sections based on feature selection methods: training and testing. Training accounts for 67% of the total records in the dataset while testing accounts for 33%. The two components are used to train and test the suggested model. Then, to make classification judgments, we'll split the training dataset over three parallel SVMs in the ensemble bagging approach and three sequential SVMs in the Ada boosting algorithm. Look at Algorithm 2.

Algorithm (2) Bagging and Ada boost with SVM Ensemble Algorithm.
Input: 5000 records of the UNSW-NB15 Dataset Output: Classification decisions(normal or attack)
Begin Steps: Step 1: Bagging Make dataset as three samples by dividing by 3 Foreach dataset sample apply SVM classifier algorithm 1. -Initialize (Xi, Yj) for all training dataset points, where X is a data vector (x1…. , xn) and Y is a class vector.

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

2. -Set the weight W vector.
3. -Allotment points of (x, y) and elicitation the hyper plane separator.
4. -Heck the hyper plane if it is provides the best separation, use it as a classifier system for the classification of the unsw nb-15 testing dataset and switch to End; otherwise, proceed to the next step.
5. -Make the hyperplan bigger.
6. -Set up the Lagrange multiplier. αi vector α1…αn.
7. -Use the classification function.
8. -Find the non-zero support vectors xi (support vectors are the points that determine the rea of hyper plan).
9. -Use the hyper plan that emerged after identifying support vectors as the classifier model to classify the unsw nb-15 testing dataset.
10. End_For
11. Make voting to return results

Step 2: Ada boosting

For i=1 to 3

For each instance of dataset

1. - Initialize equal weights to all instances of sample dataset as

$$D_i = 1\backslash N$$

2. -Train weak classifier as

   -Determine the support vector as step 2 (from 1 to 9)

3. -Compute the total error of weak classifier

$$TE = \frac{\sum_{i=1}^{N} w_i . I(y_i \neq h)}{\sum_{i=1}^{N} w_i} \quad \text{Eq(3.5)}$$

4. -Calculate the performance of weak classifier

$$Performance = \frac{1}{2} LN(\frac{1-TE}{TE}) \quad \text{Eq(3.6)}$$

5. -Reduce the weights of correctly classifed instances and increase the weights of incorrectly classifed instances as :

$$W_{new}(correct) = W_{old\ .e^{-performance}} \quad Eq(3.7)$$

$$W_{new}(correct) = W_{old\ .e^{performance}} \quad Eq(3.8)$$

6. -Normalize the weights to be =1 and create new sample based on the new weights.

End for

End for

End

V. RESULTS AND DISCUSSION

As previously stated, the goal of this article is to create a high-accuracy worm detection system. A model called CFS-chi square that combines CFS and chi2 is used to assess a subset of the original characteristics to eliminate unnecessary features and enhance classification reliability. During the classification step, the bagging and Ada boosting ensemble classifier is trained and evaluated on the UNSW NB15 dataset. The tests are run on a desktop PC with an Intel Core i3-3217U CPU and 4GB of RAM running at 1.80 GHz. True Positives (TP) (intrusion), True Negatives (TN) (normal), FP (misclassified as intrusion), false negatives (FN) (misclassified as normal), and Unknown (modern attacks) are the classification outcomes of testing. Table I shows the results of evaluating bagging

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

using SVM with four different subsets of UNSW NB-15features (All features, correlation select 30 features, chi2 select 33 features, 27 features from combining chi2 with correlation). Table II illustrates the results of using SVM to evaluate Ada boost with four distinct subsets of unsw-nb-15features (all features, correlation select 30 features, chi2 select 33 feature, 27 features from combining chi2 with correlation).

TABLE I. CLASSIFICATION RESULTS OF BAGGING WITH SVM CLASSIFIERS

Feature selection measure	TP	TN	FP	FN	Unknown
CFS&Chi2fs	928	719	1	2	0
CFS	915	717	9	9	0
Chi2	908	730	7	5	0
ALL	899	724	15	12	0

TABLE II. CLASSIFICATION RESULTS OF ADA BOOST WITH SVM CLASSIFIERS

Feature selection measure	TP	TN	FP	FN	Unknown
CFS&Chi2fs	915	718	8	9	0
CFS	892	709	22	27	0
Chi2	917	711	19	3	0
ALL	895	687	37	31	0

The Detection Rate (DR) is the ratio of the total number of intrusion records in the testing dataset to the total number of TP, DR = TP/(TP+FN+Unknown). The ratio between multiple "normal" records labeled as attacks (FP) and the entire number of "normal" records supplied in the testing dataset is the False Alarm Rate (FAR), FAR = FP/(TN+FP +Unknown). The classification accuracy, which is the ratio of the number of correctly classified patterns (TP, TN) to the total number of patterns in the testing dataset, can be used to choose the optimal model Accuracy (TP+TN)/(TP+FP+TN+FN+unknown). F-measure=(2*Recall*Precision)/(Recall+Precision),False Discovery Rate,(FDR=FP/FP+TP), Precision=TP/(TP+FP), Specificity =TN/(TN+FP). Table III and Table IV show the values for all of the metrics given.

TABLE III. METRICS TO EVALUATE ENSEMBLE BAGGING WITH SVM

Metrics	All	CFS	Chi2	CFS&Chi2fs
n.of.feature	44	30	33	27
Accuracy	0.983	0.990	0.992	0.998
DR	0.986	0.990	0.994	0.997
FAR	0.020	0.012	0.009	0.001
Precision	0.983	0.990	0.992	0.998
F-measure	0.984	0.990	0.992	0.997
Specificity	0.979	0.987	0.990	0.998
FDR	0.016	0.009	0.007	0.001

Table III highlights the results from the UNSW-NB15 dataset, which includes the ensemble bagging results using the SVM classifier. It is proposed that without feature

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

selection, the ensemble classifier is not optimal in a number of criteria. When feature selection methods are used, however, performance improves to the best possible scenario. Our proposed system achieves an accuracy is 0.983, FAR is 0.020, DR is 0.986, Precision is 0.983, F-measure is 0.984, Specificity is 0.979, and FDR is 0.016 without using feature selection methods. When applying feature selection methods, particularly when using CFS&Chi2fs, the results are optimized, where the accuracy is 0.998, FAR is 0.001, DR is 0.997, Precision is 0.998, F-measure is 0.997, Specificity is 0.998, and FDR is 0.001.

TABLE IV. METRICS TO EVALUATE ENSEMBLE ADA BOOSTING WITH SVM

Metrics	All	CFS	Chi2	CFS&Chi2fs
n. of. Feature	44	30	33	27
Accuracy	0.958	0.970	0.986	0.989
DR	0.966	0.970	0.996	0.990
FAR	0.051	0.030	0.026	0.11
Precision	0.960	0.975	0.979	0.991
F-measure	0.962	0.972	0.987	0.990
Specificity	0.948	0.969	0.973	0.998
FDR	0.039	0.024	0.020	0.008

As shown in Table IV, when the SVM classifier as base estimator in the ensemble Ada boost algorithm Without using feature selection methods, our proposed system accuracy is 0.958, FAR is 0.051, DR is 0.966, precision is 0.960, F-measure is 0.962, Specificity is 0.948, and FDR is 0.039. When using feature selection methods, the best case is 0.989, FAR is 0.013, DR is 0.992, Precision is 0.989, F-measure is 0.99, Specificity is 0.986, and FDR is 0.010. When using CFS&Chi2fs, the best case is 0.989, FAR is 0.011, DR is 0.990, Precision is 0.991, F-measure is 0.990, Specificity is 0.998, and FDR is 0.008. To better understand the benefits of the suggested methodology, we compare our proposed system to some related work. The comparison's results are shown in Table V.

TABLE V. COMPARISON OF THE PROPOSED SYSTEM WITH RELATED WORK

Method	Dataset	Feature selection	n.of.features	ACC
Ada boosting (random tree)	Diabetes Dataset	N/A	17	0.961
Bagging(SVM)	UNSW-NB15	N/A	49	0.935
Stacking (SVM)	UNSW-NB15	N/A	49	0.935
Boosting (SVM)	UNSW-NB15	N/A	49	0.882
SVM	Train and Test UNSW-NB15	N/A	44	0.85
Bossting	The payment data in taiwan	N/A	23	0.71
Bagging(SVM)	UNSW-NB15	CFS&Chi2fs	27	0.998
Ada boost(SVM)	UNSW-NB15	CFS&Chi2fs	27	0.989

Received 24/September/2021; Accepted 1/December/2021

ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

The comparison includes the classification method, the selected dataset, feature selection approaches, the number of selected features, and intrusion detection accuracy, as shown in Table V. Our suggested system outperforms Ada boosting(random tree), bagging (SVM), stacking (SVM), boosting (SVM), with an ACC of 0.998 and 0.989, respectively. When comparing our proposed system to the SVM, one can see that the ensemble method provides benefits, as the SVM is a single classifier with high variance, As a result, ensembles frequently reduce the variance component of contributing models' prediction errors, leading in a significant gain in accuracy (from 0.85 to 0.998) and a reduction in the False alarm rate.

VI. CONCLUSIONS

Data mining methods are used in this paper to detect computer worms in the network, which have a high ability to detect new varieties of computer worms automatically and reliably. The proposed strategy emphasizes the importance of Network Intrusion Detection Systems (NIDS) in identifying worm attacks, which are the most dangerous in a network and have a negative influence on resource availability. The suggested solution is more efficient due to the normalizing and discretization procedures. The correlation and chi2 algorithms are proposed as feature selection approaches to improve the proposed system's accuracy and reduce the time required. The accuracy of the Bagging classifier, which uses SVM and is aided by CFS&Chi2fs, is higher than using all features, Bagging Classifier with 30 features of CFS, or chi2 with 33 features; also, the CFS&Chi2fs has a lower false alarm rate than CFS or Chi2. The Ada boost technique, which uses SVM CFS&Chi2fs, is more accurate than using all features, Ada boost Classifier with CFS, or chi2 with 33 features; similarly, CFS&Chi2fs has a lower false alarm rate than CFS or Chi2.

REFERENCES

[1] O. c. w. d. u. ensembles, “Optimizing computer worm detection using ensembles,” 11 4 2019. [Online]. Available: https://www.hindawi.com/journals/scn/2019/4656480/. [Acesso em 3 3 2019].

[2] A. P. U. Siahaan, “Threats of Computer System and its Prevention,” *International Journal of Scientific Research in Science and Technology*, vol. 3, nº 6, pp. 448-451, 2017.

[3] S. H. a. I. A. A. Hashim, “A proposal to detect computer worms (malicious codes) using data mining classification algorithms,” *Eng. & Tech. Journal*, vol. 31, nº 2, 2013.

[4] S. O. H. Z. a. A. R. A. Al-Mamory, “IDS alarms reduction using data mining,” em *IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence*, 2008.

[5] S. H. Hashim, “Intrusion detection system based on data mining techniques to reduce false alarm rate,” *Engineering and Technology Journal* , vol. 36, pp. 110-119, 2018.

[[6] S. M. a. S. H. H. Shareef, “Proposed hybrid classifier to improve network intrusion detection system using data mining techniques,” *Engineering and Technology Journal*, vol. 38, nº 1, pp. 6-14, 2020.

[7] S. H. Hashem, “Efficiency of Svm and Pca to enhance intrusion detection system,” *Journal of Asian Scientific Research*, vol. 3, nº 4, pp. 381-395, 2013.

[8] S. H. Hashem, “Enhance network intrusion detection system by exploiting br algorithm as an optimal feature selection,” em *Threat Detection and Countermeasures in Network Security*, America , 2015, pp. 17-32.

[9] S. K. S. H. H. a. I. K. G. Majeed, “Propose hmnids hybrid multilevel network intrusion detection system,” *International Journal of Computer Science*, vol. 10, nº 5, pp. 200-208, 2013.

[10] K. K. B. Y. A. a. S. A. D. amrendra K. Singh, “Machine-Learning Based Stacked Ensemble Model for Accurate Analysis of Molecular Dynamics Simulations,” *The Journal of Physical Chemistry* , vol. 123, nº 24, pp. 5190-5198, 2019.

[11] P. Y. Taser, “Application of Bagging and Boosting Approaches Using Decision Tree-Based Algorithms in Diabetes Risk Prediction,” Multidisciplinary Digital Publishing Institute Proceedings, Turkey, 2021.

[12] H. N. a. T. V. L. Thanh, “Evaluating Effectiveness of Ensemble Classifiers When Detecting Fuzzers Attacks on the

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.22.2.8

Unsw-Nb15 Dataset," *ournal of Computer Science and Cybernetics,* vol. 36, nº 2, pp. 173-185, 2020.

[13] s. mohammad, Escritor, *Proposal algorithm to reduce false alarms in NIDS.* [Performance]. Computer Sciences Department/University of Technology, 2018.

[14] D. a. H.-B. C. Jing, "SVM based network intrusion detection for the UNSW-NB15 dataset," em *IEEE 13th international conference on ASIC (ASICON)*, china, 2019.

[15] S. K. M. K. T. M. Y. &. W. Hamori, "Ensemble learning or deep learning? Application to default risk analysis," *Journal of Risk and Financial Management,* vol. 11, nº 1, 2018.

[16] N. Saleena, "An ensemble classification system for twitter sentiment analysis," *Procedia computer science,* vol. 132, pp. 937-946, 2018.

[17] A. a. D. Z. Wosiak, "Integrating correlation-based feature selection and clustering for improved cardiovascular disease diagnosis," india, 2018.

[18] L. e. a. Ali, "Reliable Parkinson's disease detection by analyzing handwritten drawings: construction of an unbiased cascaded learning system based on feature selection and adaptive boosting model," *Ieee Access* , vol. 7, pp. 116480-116489, 2019.

[19] M. G. H. a. H. P. Shahhosseini, "Optimizing ensemble weights and hyperparameters of machine learning models for regression problems," *Machine Learning with Applications,* vol. 7, p. 100251, 2022.

[20] R. &. I. A. R. M. T. (. P. o. R. B. a. R. e. l. a. f. r. e. p. u. c. d.-l. h. r. i. B. J. o. H. 5. 1. Salam, "Potential of RT, Bagging and RS ensemble learning algorithms for reference evapotranspiration prediction using climatic data-limited humid region in Bangladesh," *ournal of Hydrology*, vol. 590, p. 125241, 2020.

[21] J. K. X. Z. X. W. L. Z. D. L. I. Liu, "Data Mining and Information Retrieval in the 21st century: A bibliographic review," *Computer Science Review,* vol. 34, p. 100193, 2019.

[22] X. X. Y. C. L. Z. W. J. A. &. Z. X. Deng, "Dynamic clustering method for imbalanced learning based on AdaBoost," *The Journal of Supercomputing,* vol. 76, nº 12, pp. 9716-9738, 2020.

Received 24/September/2021; Accepted 1/December/2021
 ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*