Evolving Systems (2021) 12:769–778
https://doi.org/10.1007/s12530-019-09317-1

ORIGINAL PAPER

Anomaly process detection using negative selection algorithm and classification techniques

Soodeh Hosseini[1,2] · **Hossein Seilani**[3]

Received: 12 July 2019 / Accepted: 2 December 2019 / Published online: 10 December 2019

Abstract
Artificial immune system is derived from the biological immune system. This system is an important method for generating detectors that include self-adaption, self- regulation and self-learning which have self/non-self-detection features. This method is used in anomaly process detection where the anomaly is non-self in the system. We present a new combining technique for anomaly process detection. This combined technique is a unification of both negative selection and classification algorithm. The main aim of the proposed techniques is to increase the accuracy in this system while decreasing its training time. In this research, CICIDS 2017 and NSL-KDD dataset with different sets of features and the same number of detectors are used. This paper presents a framework for detecting anomaly processes on a host base computer system which is established on the artificial immune system. We evaluate our technique using machine learning algorithms such as: logistic regression, random forest, decision tree and K-neighbors. Moreover, we use WEKA tool classification to perform a correlation based feature selection on the dataset.

Keywords Artificial immune system · Negative selection algorithm · Anomaly detection · Intrusion detection · Machine learning

1 Introduction

Nowadays, network usage is growing rapidly in our daily lives. Network security is becoming a great challenge as networks include a huge environmental data sharing system. This environment is versatile and popular. There are many destructive impacts that make networks a suitable target for attackers. A variety of network-based and host-based applications have been proposed to prepare many services. In addition, the increase in networked environments has led to an increase in anomaly behavioral activity with both external and internal attackers.

Since prevention processes are somewhat incomplete, monitoring is necessary for security systems. This role is assigned to an IDS. The IDS's purpose is to detect abnormal activity in real time and give an alert. We categorize two types of IDS: misuse based and anomaly based systems. Misuse based systems involve patterns which recognize analyzed data. This process is very effective in identifying known threats but it is not effective in identifying unknown threats. However, anomaly-based systems make their decisions based on system behavior, often by using statistical techniques, data mining or machine learning. Any unadaptable situation to these specifications is considered an anomaly. The following systems are capable of detecting new attacks.

Intrusion detection systems (IDS) are classified into host-based IDS (HIDS) and network based IDS (NIDS) to monitor vital operating system files (Tabatabaefar et al. 2017). This section involves a brief overview of intrusion detection approaches, classification of IDS and categories of attacks. Intrusion detection systems are presented in two types of intrusion detection methods:

✉ Soodeh Hosseini
so_hosseini@uk.ac.ir

Hossein Seilani
hosseinseilany@gmail.com

[1] Department of Computer Science, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran

[2] Mahani Mathematical Research Center, Shahid Bahonar University of Kerman, Kerman, Iran

[3] School of Computer Engineering, Bahmanyar University of Kerman, Kerman, Iran

1. Misuse detection: Misuse detection is also considered as signature-based detection. In this method, the known attack patterns are previously defined. These defined patterns function as signatures for the intrusions to be detected by the IDS. When a pair is found for the signature, a kind of warning will be shown. One of the main pros of misuse detection is its increased accuracy to identify attacks and decreased false alarms. The cons of this method is its ability to detect only known attacks.
2) Anomaly detection: In anomaly detection, a set of behaviors is derived in advance from a normal behavior which is then applied to detecting the anomaly in systems. The advantage of this method is its ability to detect new or unknown attacks and the disadvantage of this method is that it has a low detection rate and high false alarm rate. It cannot detect new anomaly.

One of the main problems of IDS in networked environments is the extensive amount of data gathered from the network and its analysis. Thus, the existing approaches of IDS focus on signature-based attacks. Many IDS contain a defense system for networked environments which uses signature-based methods for anomaly detection (Tabatabaefar et al. 2017).

The use of artificial immune system (AIS) in IDS is very important for two reasons. First, the body's immune system is highly organized to defend against external pathogens and is self-organized. Second, the current system security is unable to comply with the computer systems and their security (Brandsæter et al. 2019).

The negative selection algorithm (NSA) is one of the main AIS algorithms. This algorithm is able to detect the distinction between self and non-self naturally in the IDS area. Therefore, different AIS-based approaches are proposed to improve IDS.

Also, NSA is one of the most common Artificial Immune System models that have gained significant importance among researchers. Forrest et al. (1994) proposed a NSA which is based on the definition of self/non-self-behavior in the immune system. It is derived from the role of the T-cells negative selection in the thymus. NSA functions are based on the immune system to identify abnormal antigens or non-self-cells without affecting self-cells. It generates a series of self-patterns that are defined as normal patterns. This series can detect non-self-patterns and keep them as non-self or anomaly state. The system generates some known patterns which are known as detectors. A detector is a series of patterns to identify self-patterns from non-self-patterns. In case of miss-match between normal and oncoming pattern detectors, an anomaly is present. The compatibility and detection of abnormal attacks are the major features designated to an improved IDS. In IDS, it is clear that the system area (normal/abnormal) will vary over time. For example, computer administrators mostly change the network settings. In addition, new attacks may happen, Therefore IDS should detect normal and abnormal features in real time (Hooks et al. 2018).

In traditional IDS-based NSA, if the system is not updated, it is difficult to detect new changes in the environment. Additionally, current methods of NSA proposed for IDS require only normal records from the training class and the NSA detection step only has two classes for a tested record: normal or abnormal and no other abnormal features are provided (Wen and Li 2017).

In this method, there is a dataset that involves many samples of anomaly attacks. Attackers can change the signature of anomaly attacks to gain mastery. This is the weakness of the signature based IDS. Therefore, behavioral based detection method is implemented in the work. To meet this challenge, we used behavioral base detection methods by means of machine learning algorithms. Most of IDS methods focus on the feature of selection. The purpose of this study is to reduce input features and training time of systems in IDS. In order to do this, we investigate the performance of standard feature selection methods using Correlation-based Feature Selection. In this paper, we apply four efficient classifiers (such as logistic regression, random forest classifier, K-neighbors classifier, decision tree classifier, Gaussian) on reduced NSL-KDD datasets and CICIDS 2017 for IDS. The feature selection is then applied using standard feature selection methods and correlation-based feature selection (CFS) by using WEKA tools.

The four classified results with NSA algorithm will be computed in comparison with feature selection methods to show that our proposed model is more efficient for IDS instead of NSA algorithm. We have two purpose in this study. The first aim is to design a light weight anomaly and malware process detection using artificial immune system and NSA algorithm. The novelty of our technique is the integration of negative selection algorithm to reduce training time and setting features selection for detectors of the anomaly process. Secondly, we try to have the best feature selection to increase the accuracy in system detection and we test our algorithm by the NSL-KDD and CICIDS 2017 dataset used. The rest of this paper is organized as follows:

In Sect. 1, we briefly introduce the concepts of IDS models and NSA algorithm. Section 2 is an overview of related works and presents existing studies on AIS based IDS. Furthermore, a description of the NSA and machine learning algorithms and our experiments is stated. Section 3 presents a model overview of the results. Section 4 shows experimental results and the conclusion is presented in Sect. 5.

2 Related works

We have described the main purpose of the proposed model. Table 1 includes a review of related works and present a part of the relevant literature in recent years. Table 1 consists of two columns, Objective and Remarks, which describe the main applications of each method. Angelov et al. (2011) presents a novelty detection and use of the neuro-fuzzy system for tracking the object. Angelov (2014) presents a new less conservative and more sensitive condition for anomaly detection and so increased value of false negatives or false positives. Ugochukwu and Bennett (2018), in an article on 'An Intrusion Detection System Using Machine Learning Algorithm', investigates the application of an AIS based Intrusion IDS. The authors studied and tested two most common AIS algorithms, namely negative and clonal selection, on the NSL-KDD dataset with different sets of features and different numbers of detectors. Yang et al. (2017) proposed a model for anomaly detection which performs a high-dimensional space feature and presents a real negative selection algorithm to guide the generation of detectors. Sharma and Gupta (2017) proposed a model for intrusion detection based on NSA and J48 decision tree algorithm for the optimization of the IDS. This study applies a new hybrid model for intrusion detection which combines NSA and J48 Classification algorithms. The main aim of the proposed model was to increase accuracy and decrease false alarm rate. Zhang and Ma (2016) solicited a technique which uses I/O request for malware detection that also integrated NSA and positive selection algorithm for malware detection. Some feature selections extracted from I/O are used in NSA and positive selection algorithm (PSA). Igbe et al. (2016) submitted a model for distributed networks based on AIS. This model uses machine learning to classify traffic for detecting normal (self) and abnormal (non-self) data and, then, evaluates oncoming data based on the mentioned technique for this system. Saurabh and Verma (2016) presents an efficient proactive artificial immune system based on anomaly detection and prevention system.

3 The proposed method

In comparison to other existing IDS, the proposed model combines NSA and classification algorithm to improve training time and detection of anomaly process. For this purpose, we added the NSA algorithm in the existing IDS models. This model is shown in Fig. 1 and it includes NSL-KDD and CICIDS2017 datasets, pre-processing, NSA algorithm, classification algorithms and evaluation for improving the model. In our NSA algorithm model, two different classes of normal and anomaly dataset are used instead of normal data to train and test step. Then, the feature selection method is applied on input process data by using correlation-based feature selection (CFS) in the WEKA tools. The final result of this model will increase the rate of detection. The block diagram of our proposed model is shown in Fig. 2. It involves three phases: application of the negative selection algorithm and four classification algorithms, evaluation of the proposed model and presenting the results.

Table 1 Study of anomaly detection with NSA

References	Objective	Remarks
Angelov et al. (2011)	Autonomous novelty detection	Presents new approaches to both the problem of novelty detection and object tracking in video streams. it does not require a user- or problem-specific threshold to be pre-defined
Angelov (2014)	Anomaly detection based on eccentricity analysis	propose a new condition for anomaly detection. use data analytics. for several types of similarity measures (such as Euclidean, cosine, Mahalanobis)
Ugochukwu and Bennett (2018)	Intrusion detection system	Application of the artificial immune system for intrusion detection investigates the application of an AIS based intrusion IDS
Yang et al. (2017)	Anomaly detection based on NSA	Proposed a model for anomaly detection which performs in high-dimensional space features and presents a real negative selection algorithm to guide the generation of detectors
Sharma et al. (2017)	Intrusion detection based on NSA	This study presents a new hybrid model for intrusion detection which combines NSA and j48 classification algorithms
Zhang et al. (2016)	Malware detection by NSA	Proposed a technique which uses i/o requests for malware detection: integrated NSA and positive selection algorithm for malware detection
Igbe Ihab et al. (2016)	Network intrusion detection by NSA	Proposed a model for distributed networks based on AIS
Saurabh et al. (2016)	Anomaly Detection by NSA	An efficient proactive artificial immune system based on anomaly detection and prevention.

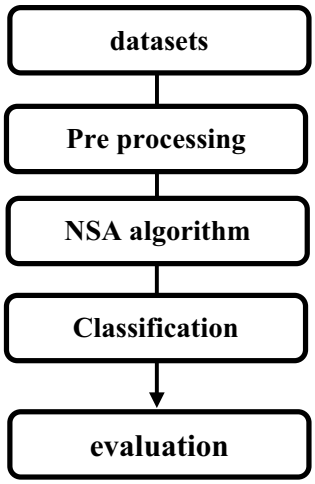

Fig. 1 Proposed model

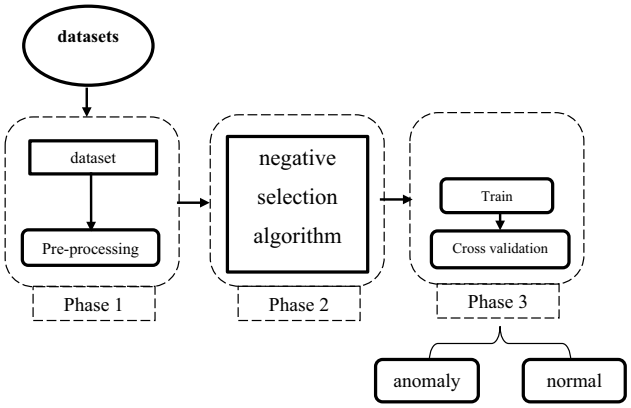

Fig. 2 Block diagram of proposed model

The NSA has two separate steps: First, detectors are randomly defined and produced in one step and then assigned to self-actions. Next, the system compares new input data and controls the system. It checks data using existing detectors and if any new data is higher than the specified threshold value, it rejects the data and activates a detector and alerts system administrator.

NSA has been applied to monitor changes in the system. This model is presented for the detection of anomaly processing which, in turn, helps increase detection accuracy and reduces system training time.

3.1 Dataset

The first step of each learning process is data gathering. Appropriate data is helpful to reach better results and design an efficient framework. Results of the presented framework are based on two distinct datasets:

The first dataset is NSL-KDD. NSL-KDD is an improved dataset derived from KDDCUP'99, which solves several problems of KDDCUP'99 and prepares a new dataset with selected records from KDD dataset which show that previous problems do not exist anymore. In this study we used the last version of NSL-KDD dataset. The NSL-KDD Dataset is a famous dataset in the field of IDS. This dataset includes a standard set of data with 41 features and 4 different attack classes. It also contains 4 types of attack class: DOS, R2L, U2R, PROB. Each record in KDD dataset has 41 parameters (containing 1 class Label, 3 symbolic and 38 continuous fields) which are separated by columns.

An example of this dataset record is shown as below:

Normal process																			Normal								
0	0	0	0	136		1	0	0	1	1	0.01		0.06		0	255	1	0	0.06		0	0	0	0	1	1	

Anomaly process																											
1	0	0	0	1	1	0	0	0	0	1	0	0	134		86	0.61		0.04		0.61		0.02		0	0	0	0

The second dataset is a dataset derived from CICIDS 2017 dataset, presented in 2019, which have been used in order to effect the feature selection method on our target model. This dataset includes 83 features and 15 defined class labels that include 14 attack and 1 normal class. Before any operation is conducted on these datasets, data is transformed into numerical and then normalize data (Panigrahi and Borah 2017).

3.2 Pre-processing

Since there is a lot of raw data in the incoming traffic, selection of a special set of features that can increase detection, especially for algorithms that are sensitive to the number of features, is needed. "Deciding upon the right set of features" is difficult and time consuming therefore we use the WEKA tool. WEKA is an open source program made for programming in JAVA language and was made to implement a variety of learning machines and data mining models. It involves 76 classification algorithms, 49 data pre-processing, 10 search algorithms for feature selection and 15 attribute evaluators (Brown et al. 2016). All features have not been linearly correlated and their value has nearly reached zero. Therefore, using Pandas library, we got the value of correlation, and deleted the values near zero to reduce the noise. These 20 features have the highest degree of correlation on the prediction class. The correlation is one of the most commonly used statistical concepts. In this article we show how to calculate the correlation by using the panda python library. The term "correlation" refers to the association or relationship between quantities. We show the correlation formula by using two variables of X and Y:

```
=== Attribute Selection on all input data ===
Attribute ranking.
Correlation Ranking Filter
Ranked attributes:
 1.  0.70701  29 SameSrvRate
 2.  0.64378  39 DstHostSrvSerrorRate
 3.  0.64074  38 DstHostSerrorRate
 4.  0.63916  25 SerrorRate
 5.  0.63874  26 SrvSerrorRate
 6.  0.62058  34 DstHostSameSrvRate
 7.  0.60807   4 Flag
 8.  0.60704  33 DstHostSrvCount
 9.  0.56885  12 LoggedIn
10.  0.50554  23 Count
11.  0.37092  32 DstHostCount
12.  0.26442   3 Service
13.  0.18732  36 DstHostSameSrcPortRate
14.  0.18679  40 DstHostRerrorRate
15.  0.18646  27 RerrorRate
16.  0.18629  41 DstHostSrvRerrorRate
17.  0.18506  28 SrvRerrorRate
18.  0.17743  35 DstHostDiffSrvRate
19.  0.17042   2 ProtocolType
20.  0.16664  31 SrvDiffHostRate
```

Fig. 4 WEKA feature selection result

In formula 1, observations of raw data are made by measuring standard deviations. x_i, y_i is an example of the vector x, y, and X_j, Y_j is the mean of the vector x, y.

$$P_{x,y} = \frac{\sum (x_i - X_j)(Y_i - Y_j)}{\sqrt{\sum (x_i - X_j)^2 \sum (Y_i - Y_j)^2}} \quad (1)$$

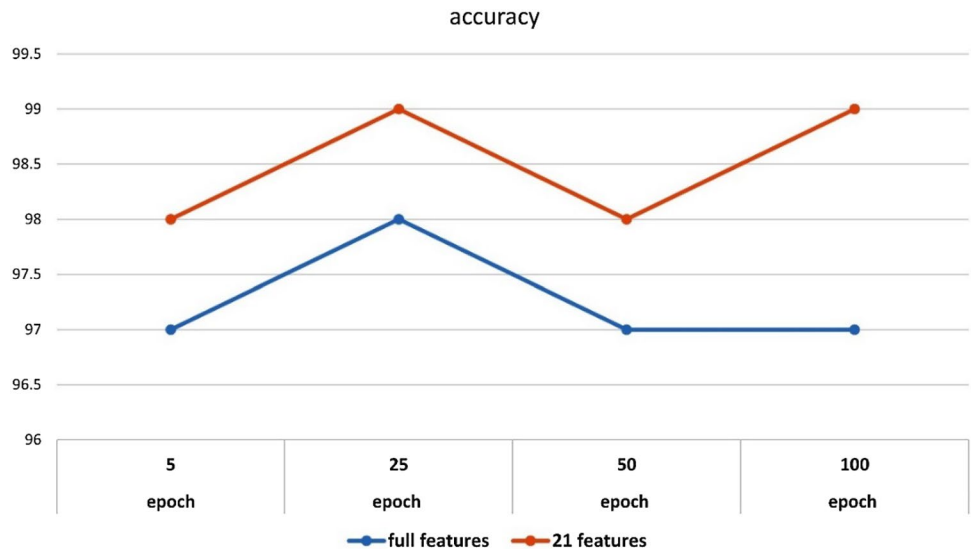

Fig. 3 Block diagram of proposed model

In Fig. 3, the number of epochs is shown. As expected, using the negative selection algorithm to reduce the number of feature selections has reduced the training time and stabled the accuracy.

These features are selected based on the information gained from WEKA tool (Meena and Choudhary 2017). A snapshot of feature selection is presented in Fig. 4.

The experiment has been measured using the correlation between it and the class method in others to remove the irrelevant features from the datasets. Using the correlation technique, 20 features (29, 39, 38, 25, 26, 34, 4, 33, 12, 23, 32, 3, 36, 40, 27, 41, 28, 35, 2, 31) are selected from 41 features to reduce the time needed for training and detection.

3.3 Data wrangling

The dataset has null values that are called NAN. So, we cleaned up the duplicates, blanks and simple errors data using the data wrangling method for cleaning messes and complex dataset for analysis and easy access. The data is finally stored into one single dataset. Data wrangling methods ensure that we have unified the data. The Numpy and Pandas library are two famous and powerful ways for data analysis and performing matrix relationships, especially in machine learning operations. We have also used Numpy library to find and clean the NAN data from dataset. Afterwards, they are removed or replaced in value from the dataset with Pandas library. Grouping, filtering and choosing appropriate data will increase accuracy in the proposed model and quicken the process of decision making in the learning machine.

3.4 NSA algorithm

The NSA algorithm is based on self-set (normal) and non-self-set (anomaly) for behavioral detection (Pharate et al. 2015). The main function of this system is to recognize normal and anomaly process (Xu et al. 2019). It has a set of self-patterns for identifying non-self or anomaly. These patterns are known as detectors. In this paper, we use the cosine similarity to measure similarity:

Cosine similarity formula:

$$Sim(A.B) = COS(\theta) = \frac{A.B}{|A||B|} \tag{2}$$

In formula 2, values A, B are two vectors, and θ is the angle between these two vectors. The training dataset includes anomaly and normal process data. We train the system with this dataset. Also, a normal process data set is considered as a detector which is not included in the training dataset. In general, we are comparing each row of training data with this detector and obtain the hamming space. If the hamming space is less than the threshold, then the row will be rejected, otherwise we will save it for the training phase. The following pseudo code 1 shows the proposed algorithm method:

```
Assume detector is a positive set of system

//NSA phase
For each value in columns:
        If simulate (value[i], detector[i]) >=threshold
Return 0;
Else
Dataset. Insert (value[i]);

// training phase
Dataset=Read (Dataset)
Do Classification (Logistic Regression, Random Forest Classifier, K-Neighbors Classifier, Decision Tree
Classifier)

// classification phase
For each row in dataset:
if (row== detector) then:
        Print (row data is non-self)
Else
        Print (data is self)
```

Table 2 Logistic regression with 20 featuers

NSL-KDD dataset	Precision	Recall	f1-score	Accuracy
Without NSA	84	83	88	88
With NSA	95	95	95	95

Table 3 Random forest with 20 features

NSL-KDD dataset	Precision	Recall	f1-score	Accuracy
Without NSA	98	98	97	97
With NSA	99	99	99	99

Table 4 K-neighbors with 20 features

NSL-KDD dataset	Precision	Recall	f1-score	Accuracy
Without NSA	94	94	96	95
With NSA	95	95	95	97

Table 5 Decisiontree with 20 features

NSL-KDD dataset	Precision	Recall	f1-score	Accuracy
Without NSA	97	97	96	97
With NSA	99	99	99	98

3.5 Evaluation

Performance can be measured by using the following formulas:

TP: True Positive (): Anomaly attack that is correctly classified as Anomaly attack.
TN: True Negative (): Normal data that is correctly classified as normal data.
FN: False Negative (): Anomaly attack incorrectly classified as normal data.
FP: False positive (): Normal data incorrectly classified as malicious attack (Johny et al. 2017).
a) Accuracy: It is the proportional correction of TP and TN classification over the total number of classifications and can be calculated by the formula (Pharate et al. 2015):

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN}$$

b) Precision: It is known as the possibility of a positive prediction that is being correct.

$$\text{Precision} = \frac{TP}{TP + FP}$$

c) Recall: recall is the number of correct results divided by the number of results that should have been returned.

$$\text{Recall} = \frac{TP}{TP + FN}$$

F1 score: calculating a mean of precision and recall, in a way that seek a balance between precision and recall.

$$\text{f-score} = 2 * \frac{(\text{Precision*recall})}{\text{Precision} + \text{recal})}$$

Table 6 Comparison of evaluation performance with NSA

Algorithm	Evaluated with NSA							
	NSL-KDD dataset				CICIDS 2017 dataset			
	Precision	Recall	f1-score	Accuracy	Precision	Recall	f1-score	Accuracy
Logistic regression	95	95	95	95	94	94	94	94
Random forest	99	99	99	99	97	97	97	97
K-neighbors	95	95	95	97	96	94	94	94
Decision tree	99	99	99	98	97	97	97	98

Table 7 Comparison of evaluation performance without NSA

Algorithm	Evaluated without NSA							
	NSL-KDD dataset				CICIDS 2017 dataset			
	Precision	Recall	f1-score	Accuracy	Precision	Recall	f1-score	Accuracy
Logistic regression	84	83	88	88	82	82	87	86
Random forest	98	98	97	97	97	97	97	97
K-neighbors	94	94	96	95	93	93	92	93
Decision tree	97	97	96	97	96	97	96	96

Table 8 Evaluation of the result of accuracy without NSA

Algorithms	Logistic regression	Random forest	K-neighbors	Decision tree
Accuracy without NSA (NSL-KDD)	88	97	95	97
Accuracy without NSA (CIC-IDS2017)	86	97	93	96

4 Experimental results

This section includes the experimentation of the results that have also been compared with the other five classifiers. The NSL-KDD and CICIDS 2017 datasets is used for the experiment. It involves 292,300 processes. Python programming language is used to develop the proposed IDS. Evaluation to detect the attacks using the proposed model gives more reasonable results than the five algorithms. In this evaluation, 5 k-fold cross validation is used to avoid overtraining and overfitting. Results of the evaluation differed from k-fold 0 to 5. The amount of k-fold did not change from 5 to 10. Therefore, the amount of k-fold 5 has been used for the evaluation process. The proposed approach (NSA + machine learning algorithms) is more effective than the machine learning algorithms. The proposed model is a combination of negative selection algorithms and four classification algorithms, shown in the list, which enhance the efficiency of IDS:

1. Logistic regression.
2. Random forest classifier.
3. K-neighbors classifier.
4. Decision tree classifier.

Table 2 shows the effects of a few selected features on the logistic regression. Results of the evaluation performance increased when the number of features decreased. Cutting the number of features in half resulted in increased levels of accuracy value. Comparison of evaluation performance on the logistic regression algorithm is shown in Table 2.

Table 3 shows the effects of a few selected features on the random forest. Results of the evaluation performance increased when the number of features decreased. Cutting the number of features in half resulted in increased levels of accuracy value. Comparison of evaluation performance on the random forest algorithm is shown in Table 3.

Table 4 shows the effects of a few selected features on the K-neighbors. Results of the evaluation performance increased when the number of features decreased. Cutting the number of features in half resulted in increased levels of

Table 9 Evaluation of the result of accuracy with NSA

Algorithms	Logistic regression	Random forest	K-neighbors	Decision tree
Accuracy with NSA (NSL-KDD)	95	99	97	98
Accuracy with NSA (CICIDS2017)	94	97	94	98

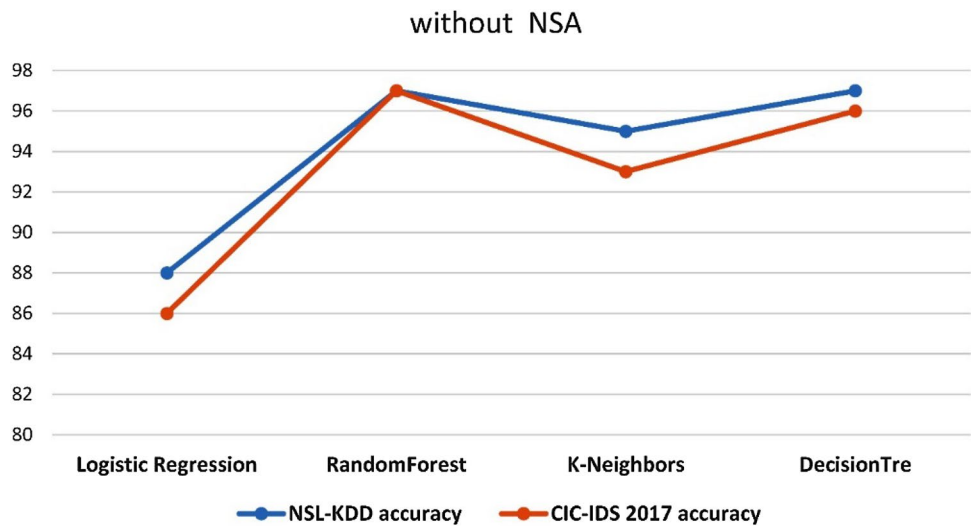

Fig. 5 Accuracy compare between algorithms for two datasets, without NSA

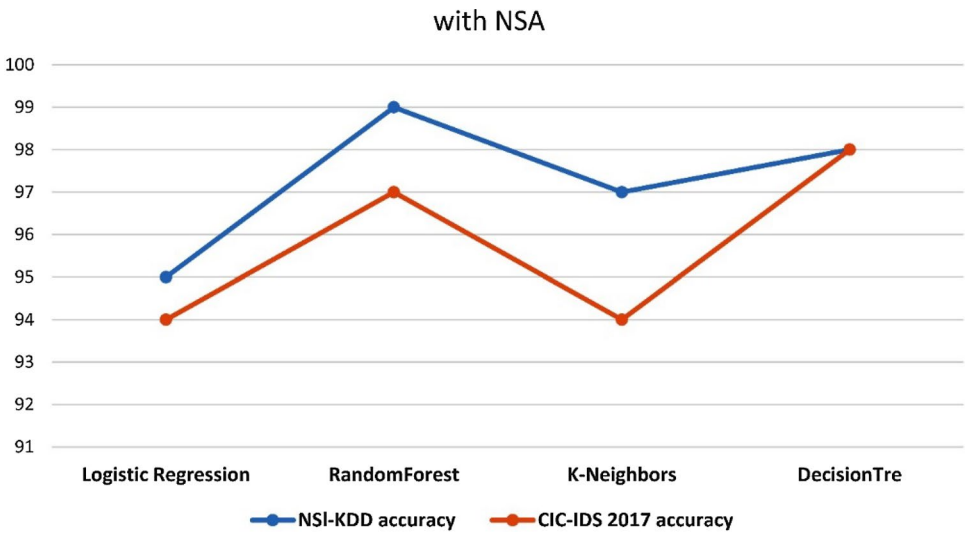

Fig. 6 Accuracy compare between algorithms for two datasets, with NSA

accuracy value. Comparison of evaluation performance on the K-neighbors algorithm is shown in Table 4.

Table 5 shows the effects of a few selected features on the decision tree. Results of the evaluation performance increased when the number of features decreased. Cutting the number of features in half resulted in increased levels of accuracy value. Comparison of evaluation performance on the decision tree algorithm is shown in Table 5.

Table 6 shows the Comparison of evaluation performance on all the machine learnings algorithm with NSA method on two datasets.

Table 7 shows the comparison of evaluation performance on all the machine learnings algorithm without NSA method on two datasets.

Table 8 shows the effects of accuracy on all the algorithms. The comparison of evaluation performance on the all algorithms is shown in Table 8 without NSA.

Table 9 shows the effects of accuracy on all the algorithms. The comparison of evaluation performance on the all algorithms is shown in Table 9 with NSA.

Figures 5 and 6 are created based on Tables 8 and 9; the X-axis displays the machine learning classifiers while the Y-axis display the percentage of accuracy. Figure 5 compares the performance of all algorithms without NSA and Fig. 6 with NSA algorithm on two datasets. All algorithms with NSA are more accurate and more effective than the algorithm without NSA algorithm approach.

5 Conclusion

In this paper, a new model is presented to detect anomaly processes in computer systems. In this model, the AIS-based NSA algorithm is used to reduce system training time by maintaining the accuracy in detection. We used the latest NSL-KDD as well as the CICIDS 2017 datasets which are commonly used in IDS. Due to the lack of NSA in IDS to detect anomaly processes, we are encouraged to provide a model which is able to detect anomaly processes. In this case, there is no need to use signature databases. This model uses a behavioral-based detection method. The study is conducted using 41 attributes of NSL-DATASET, among which 20 attributes are selected by using the correlation method of WEKA tool which has the highest degree of correlation in the prediction class for eliminating the noises in data. The system is trained by using the 5 k-fold method. Consequently, using machine learning algorithms and two datasets with considering the NSA approaches, we are reached more accuracy about 99% in a runtime range which obtained 27.36 s. On the other hand, without using the NSA method, accuracy value time reached to 33.74 s.

References

Angelov P (2014) Anomaly detection based on eccentricity analysis. In: 2014 IEEE symposium on evolving and autonomous learning systems (EALS). pp 1–8. IEEE, New York

Angelov P, Sadeghi-Tehran P, Ramezani R (2011) An approach to automatic real-time novelty detection, object identification, and tracking in video streams based on recursive density estimation and evolving Takagi-Sugeno fuzzy systems. Int J Intell Syst 26(3):189–205

Brandsæter A, Vanem E, Glad IK (2019) Efficient on-line anomaly detection for ship systems in operation. Expert Syst Appl 121(1):418–437

Brown J, Anwar M, Dozier G (2016) Intrusion detection using a multiple-detector set artificial immune system. In: 17th international conference on information reuse and integration (IRI). pp 283–286

Forrest S, Perelson AS, Allen L, Cherukuri R (1994) Self-Nonself discrimination in a computer. In: Proc. 1994 IEEE symp. on security and privacy, pp 202–212

Hooks D, Yuan X, Roy K, Esterline A, Hernandez J (2018) Applying artificial immune system for intrusion detection. In: 2018 IEEE fourth international conference on big data computing service and applications (big data service), Bamberg, pp 287–292

Igbe O, Darwish I, Saadawi T (2016) Distributed network intrusion detection systems: an artificial immune system approach. In: IEEE first international conference in connected health: applications, systems and engineering technologies (CHASE). pp 101–106

Johny D, Haripriya P, Anju J (2017) Negative selection algorithm: a survey. Int J Sci Eng Technol Res 6

Meena G, Choudhary RR (2017) A review paper on IDS classification using KDD 99 and NSL KDD dataset in WEKA. In: International conference on computer, communications and electronics (Comptelix). pp 553–558

Panigrahi R, Borah S (2018) A detailed analysis of CICIDS2017 dataset for designing intrusion detection systems. Int J Eng Technol 7(3.24):479–482

Pharate A, Bhat H, Shilimkar V, Mhetre N (2015) Classification of intrusion detection system. Int J Comput Appl 118:23–26

Saurabh P, Verma B (2016) An efficient proactive artificial immune system based anomaly detection and prevention system. Expert Syst Appl 60:311–320

Sharma S, Gupta RK (2017) A model for intrusion detection based on negative selection algorithm and J48 decision tree. Int J Res Appl Sci Eng Technol 5:1–7

Tabatabaefar M, Miriestahbanati M, Grégoire J-C (2017) Network intrusion detection through artificial immune system. In: 2017 annual IEEE international on systems conference (SysCon). pp. 1–6

Ugochukwu CJ, Bennett E (2018) An Intrusion detection system using machine learning algorithm. Int J Comput Sci Math Theory 4:2545–5699

Wen C, Tao L (2017) Parameter analysis of negative selection algorithm. Inf Sci 420:218–234

Xu K, Xia M, Mu X, Wang Y, Cao N (2019) EnsembleLens: ensemble-based visual exploration of anomaly detection algorithms with multidimensional data. IEEE Trans Visual Comput Graphics 25:109–119

Yang T, Chen W, Li T (2017) A real negative selection algorithm with evolutionary preference for anomaly detection. Open Phys 15:121–134

Zhang F, Ma Y (2016) Integrated Negative Selection Algorithm and Positive Selection Algorithm for malware detection. In: International conference on informatics and computing (PIC). pp 605–609