ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

Analysing Anomaly Process Detection using Classification Methods and Negative Selection Algorithms

Sakshi Sharma[1*], Natasha Dutta[2]

Desktop Infrastructure Analyst, Midland Credit Management, San Diego, CA[1]

Project Engineer, Online Micro Services, India[2]

ABSTRACT: The primary objective of the study of anomaly process identification using negative selection algorithms and classification techniques is to improve the ability to identify deviations from expected patterns in complex data sets. Inspired by biological immune systems, negative selection algorithms provide a novel approach to anomaly discovery by efficiently distinguishing between normal and abnormal data sets. When combined with various categorisation techniques, these methods can improve anomaly detection systems' accuracy and robustness. This abstract explores the use of negative selection algorithms in conjunction with both traditional and advanced classification methods to optimise anomaly detection processes. By using these connected approaches, the project hopes to address problems such as false positives, detection delay, and adaptability to different data contexts. The findings suggest that the synergy of negative selection algorithms and classification techniques can lead to more precise and reliable detection of anomalies, providing valuable insights for applications across cybersecurity, finance, healthcare, and other critical fields.

KEYWORDS: Artificial immune system, security, negative selection algorithm, anomaly detection, Fraud detection, Neural networks, Data analysis, Outlier detection.

I. INTRODUCTION

Anomaly detection is a technique used to identify data points or patterns that significantly deviate from the expected norm. It is a crucial tool for uncovering rare and often critical events that may indicate underlying issues or potential threats. The process involves analyzing a dataset to establish what constitutes "normal" behavior, and then monitoring the data to detect instances that differ from this established norm [1].

Anomaly detection is widely used in various domains. In cybersecurity, it helps to identify unusual activities that could signify a cyber attack. In finance, it is employed to spot irregular transactions that might indicate fraudulent activities. In healthcare, it can detect unusual patient data that may point to a potential health issue. In manufacturing, it helps in monitoring equipment performance to predict maintenance needs before failures occur.

The effectiveness of anomaly detection hinges on the ability to define what is "normal" and to distinguish it from deviations. Techniques used for anomaly detection range from statistical methods and machine learning algorithms to more complex approaches like neural networks. The choice of technique often depends on the nature of the data and the specific application, but the goal remains the same: to ensure timely identification of anomalies to mitigate risks and improve outcomes [2-5].

1.1 Importance of Anomaly Detection

Anomaly detection is a pivotal technique in the realm of data analysis due to its significant role in identifying patterns that deviate from the norm, which can signal underlying issues or potential threats. Its importance spans various sectors, including cybersecurity, where it is essential for detecting unusual network behaviour that may indicate a cyber attack or data breach. In financial sectors, anomaly detection helps in spotting irregular transactions that could signify fraudulent activity or financial anomalies, thereby protecting assets and maintaining integrity. In healthcare, it is crucial

ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

to monitor patient data to detect anomalies that may indicate serious health conditions or deteriorations, leading to timely medical interventions. Additionally, in manufacturing and industrial processes, anomaly detection assists in identifying equipment malfunctions or inefficiencies before they result in costly breakdowns or operational downtimes. The ability to promptly identify and address these deviations not only helps in mitigating risks but also enhances the overall efficiency and security of systems. As data volumes grow and systems become increasingly complex, the importance of robust anomaly detection methods becomes even more pronounced, making it a cornerstone of modern data-driven decision-making and risk management strategies.

1.2 Comparative Analysis of Negative Selection Algorithms and Traditional Anomaly Detection Methods

A comparative analysis of negative selection algorithms and traditional anomaly detection methods provides insight into the strengths and limitations of these approaches in identifying unusual patterns within data. Traditional anomaly detection methods, such as statistical techniques and distance-based approaches, rely on predefined models of normal behavior and are often based on assumptions about data distributions. Statistical methods, for example, typically use measures like mean and variance to determine thresholds for anomaly detection, which can be effective in controlled environments but may struggle with high-dimensional or complex data. Distance-based approaches, such as k-nearest neighbors, measure the distance between data points to identify outliers, which can be computationally intensive and sensitive to the choice of distance metrics [6]. Negative selection algorithms, inspired by the immune system, offer a different paradigm by focusing on identifying anomalies based on their deviation from a set of known normal patterns. These algorithms generate a set of "detectors" that represent the normal data, and any data point that does not match these detectors is considered anomalous. This approach can be more adaptive to varying data patterns and does not rely on strict assumptions about data distribution. Negative selection algorithms are particularly effective in scenarios where the normal behavior is complex or not well-defined, as they can dynamically adjust to new patterns and anomalies. The comparative analysis reveals that while traditional methods are often simpler and easier to implement, they may lack flexibility and robustness in dynamic or high-dimensional environments. Negative selection algorithms, on the other hand, offer a more adaptable and potentially more accurate approach to anomaly detection but can be more computationally demanding and complex to implement. Ultimately, the choice between these methods depends on the specific requirements of the application, including the nature of the data, the computational resources available, and the desired accuracy of anomaly detection.

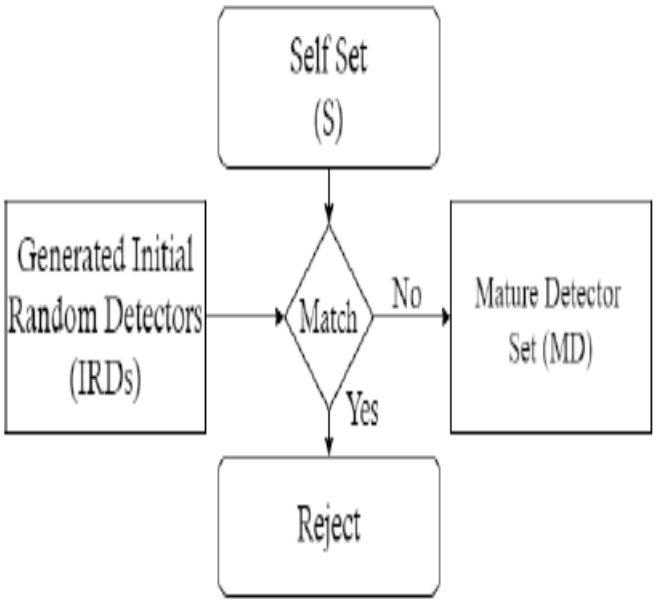

Figure 1: Generation of valid detector set.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

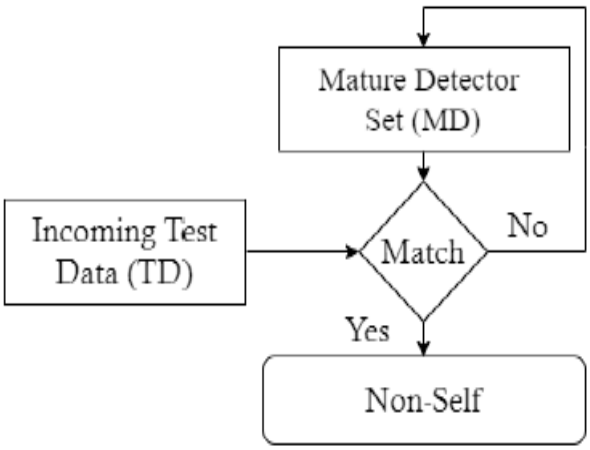

Figure 2: Detection of non-self

1.3 Pre-Processing and Dimensionality Reduction

The pre-processing transforms the training as well as test dataset into significant form for efficient processing. Normalization process normalizes pre-processed data in the range of 0 to 1. Dimensionality reduction aids in selecting better-qualified detectors by reducing the search space for the detector. NSL-KDD is most popular, universally acceptable, and recognized dataset [15]. Therefore, in this research work for experimentation, NSL-KDD dataset has been used. This dataset has 1,48,517 records and each record represents a Transmission Control Protocol/Internet Protocol (TCP/IP) link that consists of 41 features plus a "normal" or "attack" mark. This huge number of dimensions makes it very difficult and time-consuming for computation. To resolve this dimensionality challenge, this work implements the hybrid dimensionality reduction technique by using Column Standardized Normalization followed by Stacked autoencoders (SAEs) and random forest. Column Standardized Normalization is used to normalize the main network components in the range [0-1]. As shown in Figure 3, hybrid dimensionality reduction algorithm works in two phases. In first phase, the deep features are extracted by SAEs. To reduce the features further, random forest feature selection method has been applied which results in most critical features.

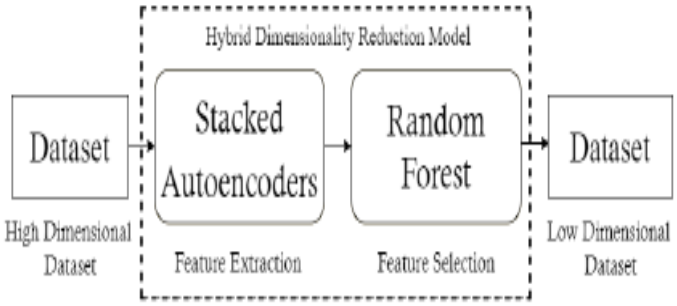

Figure 3: Flow diagram of hybrid dimensionality reduction technique.

II. LITERATURE REVIEW

The NSA is described by a number of researchers. Its most commonly used representations are binary and real-valued. As any data is ultimately translated to binary bits, therefore, the study focuses on binary representation and AIS coding scheme one of the most widely accepted. Due to string length limitations of binary representation many of the immunity features cannot be expressed. Binary representation is sufficient to depict categorical attributes. Because of these reasons, much of the work in the NSA uses binary representation along with different affinity measures, such as r-

ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

contiguous bit matching [8], r-chunk matching, Hamming distance [18], and Rogers and Tanimoto (R and T) matching [13]. Forrest et al. [8] presented the first binary string theory since it was a finite space that made problem space analysis simple. The NSA splits the 32-bit string into eight substrings, including antigen and antibody. The r-contiguous bit matching technique was used to generate the detectors. To evaluate the performance of their proposed model, they conducted three set of experiments;
1. Using random binary strings.
2. On SPARC intrusions generated by compiling C programs.
3. On COM files infected with computer viruses in Disk Operating System (DOS) environment.

In their experiments the r-contiguous value varied from 1 to 13 and total number of detectors varies from 50 to 100. They demonstrated that their proposed system can detect 50% to 85% of the changes occur in the system. This work was very initial effort to apply the AIS in intrusion detection.

The most works were limited to the binary representation of data and detectors. Subsequently, numerous attempts have been made using different methods to build an effective detector generation algorithm. However, citing the weakness of the NSA algorithm, Gonzalez et al. [10] suggested a Real-Valued Negative Selection (RNS) algorithm. The RNS algorithm represents self, detectors and non-self with real-values to resolve the inadequacies of binary representations. The present work will also use real valued NSA. Detectors in the RNS algorithm were n-dimensional vectors with a radius r in the hypersphere. The Euclidean matching function was used to match the detector with any input pattern. Gonzalez et al. [9] introduced a randomised, Real-Value, Negative Selection (RRNS) algorithm. This algorithm calculates the number of detectors needed by using the Monte Carlo method to cover non-self-space. They suggested that the number of holes and unaddressed spaces were effectively reduced by using smaller radius detectors, as it requires fewer computations. Stibor et al. [23], compared real-valued positive and negative selection algorithms with two other statistical anomaly detection algorithms Support Vector Machine (SVM) and Parzen-Window. The experiment was conducted on high dimensional Knowledge Discovery in Dataset (KDD) dataset and the investigations revealed that the NSA with variables sized detectors is not competent to real valued positive selection algorithm and statistical anomaly detection techniques on KDD dataset. Balachandran et al. [6] proposed a system for the generation of multi-form detectors in real-valued NSAs. They extended real-valued NSA by using multi shaped detectors (sphere, rectangle or ellipse) to cover two-dimensional non-self-spaces. Subsequently, Ji and Dasgupta [11, 12] suggested a new real-valued NSA that would produce variable size detectors. Detectors were represented as circles in two-dimensional spaces, and the radii of these circles were variable. On the other hand, Ji and Dasgupta [11] expanded the RNS algorithm with the variable detector radius. This work successfully demonstrated an increase in detection accuracy and protected non-self-space with fewer detectors. They conclude that smaller radius detectors decreased the number of holes and unaddressed spaces.

III. DETECTOR SELECTION

In this step, the mature detectors are selected based on the three algorithms namely NSA_ED, NSA_CD and NSA_PD. As shown in Algorithm (1) and Figure 4, the Initial Random Detectors (IRD) are matured using three algorithms separately. IRD set containing detectors (d1, d2, ……dn) are n randomly created vectors. Each IRD is matched against the self-set of training data instances by forming a similarity measure matrix. From the similarity measure matrix, for any particular IRD, the closely related affinity value is selected among all the data instances. This indicates that these detectors are how closely related to any data instance. From this closely related list, the detector that has the highest affinity value is selected as a mature detector. All other detectors are dropped out because these detectors are closely related to self-instance. This process is repeated until the desired number of mature detectors are selected from IRD.

ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

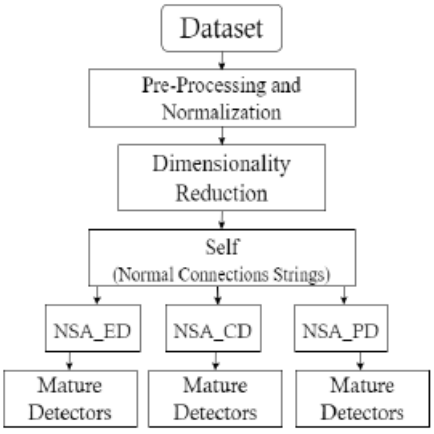

Figure 4: Flow chart of mature detectors selection algorithm.

Although there is a chance that the selected detector is not so mature for every iteration, this problem is solved by self-tuning the mature detectors. Self-tuning of the detector is done based on the ranking value. The detectors that are highly apart from all the instances of training data are marked as highly ranked. Based on this, the highly ranked detectors are selected, and the low ranked detectors are deleted from the list. No doubt this process consumes more time and also the detector rejection rate is also high, but this process increases the power of mature detectors which helps to increase the detection rate.

IV. INTRUSION (NON-SELF) DETECTION

In the third step, mature detectors are used to detect the anomalous (non-self) instance in the data set. As shown in Algorithm (2) and Figure 5, the mature detectors are matched with all test data instance by forming a similarity measure matrix using three similarity measures ED, CD and PD separately. From the similarity matrix, for any particular data instance, the affinity of the all other detectors are calculated from one particular closely related detector. This indicates how far the other detectors are from that particular detector which is matched with that particular test instance. It indicates that this data point may be anomalous, but the final decision is not made based on this single matched detector. All the other affinity values are compared with the binding threshold value. Binding threshold is the affinity value between test data instance and the mature detector. If the compared value of affinity is less than the binding threshold value, then raise the temporary alert alarm. Count all the temporary alert alarm for that particular data instance and compare with the matching threshold. Matching threshold indicates the total number of detectors matched with particular test instance. If the number of alert alarms is more than the matching threshold, then raise the final alarm for non-self; otherwise, data instance is self. This procedure is repeated for all the instances in the test data set. The decisions of NSA with different measures are calculated individually, as shown in Figure 6.

Algorithm 2: Testing phase based on NSA_ED, NSA_CD and NSA_PD
Input: TD, MD, B_T, M_T, where TD=test set (t1, t2,……. tm); dxi=detector; IRD=set of initial random detectors; MD=mature detector set (d1, d2,……. dn); SMTnxm=similarity matrix having n detectors and m data points; B_T = binding threshold; M_T= Matching threshold; TAA=Temporary alert alarm;
Step: 1 Start
Step: 2 For all instances of TDi=1:m
Step: 3 Match all the instance of MD with TD by using similarity measures Step: 4 Calculate the SMTnxm between MD and TD
Step: 5 Calculate how far the other detectors are as compare to one closely related detector to one data instance ti.
Step: 6 For data instance ti, if SMTi,j<B_T, then increment TAA
Step: 7 Repeat the step 6 for all detectors
Step: 8 If TAA>M_T, then raise the alarm for non self Otherwise, data instance is self

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

Step: 9 Repeat the step 5 to 8 for all data instances in TD
Output: Data instance having intrusion

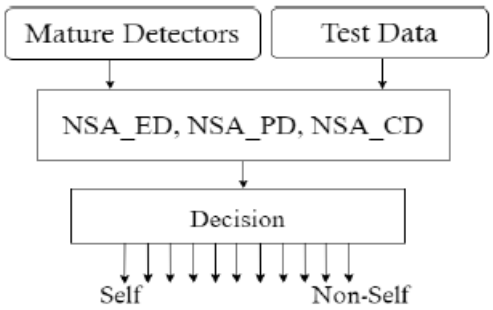

Figure 5: Flow chart of testing of NSA_ED, NSA_CD and NSA_PD algorithms.

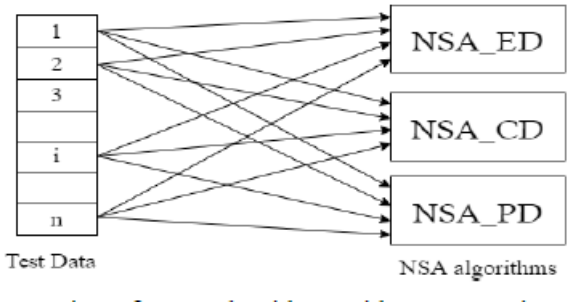

Figure 6: Testing of NSA algorithms with same test instances of NSL-KDD dataset.

V. ENSEMBLE VOTING

The final testing is carried out in this step of the proposed technique, as shown in Algorithm (3) and Figure 7. The decision of NSA based on different measures has been passed to the ensemble voting algorithm to make the final decision about any test instance. Based on the majority votes, the final decision is made whether the tested data instance is normal or anomalous. The proposed technique

Ensemble Voting based Intrusion Detection Technique using Negative Selection Algorithm 155 increases the learning rate by rectifying the false decisions made by NSA_ED, NSA_CD and NSA_PD, when runs independently. The proposed NSA_EV improves the performance in term of DR and FAR, by combining the prediction power of the different NSA algorithms.

4. Experimental Results and Analysis In this paper, all the experiments are carried out on the NSL-KDD dataset on system having Nvidia Graphic Processing Unit version GeForce GTX 1080 with 2560 compute Unified Device Architecture (CUDA) cores and a graphics clocks of 1607 MHz. Python Anaconda tool is used to process the experimental analysis part of the work. NSL-KDD dataset has been commonly used as a reference dataset for identification of anomalies in computer security problems. For experimentation, the test set used in this work consists of 5000 randomly selected undetected data, which includes both self and non-self-data. All the results are computed using the average of 40 runs in the same configuration. The following performance metrics have been used for performance evaluation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

Algorithm 3: Ensemble Voting Algorithm for intrusion detection
Input: A_L, PL_ED, PL_PD, PL_CD; A_L=Actual labels set; PL_ED=predicted labels set of NSA_ED, PL_PD=predicted labels set of NSA_PD; PL_CD=predicted labels set of NSA_CD
Step: 1 Start
Start: 2 Generate voting vector V(i) based on the votes from {PL_ED(i), PL_PD(i), PL_CD(i)}
Start: 3 If V(i)>= two for non_self votes, then raise the alarm as non self
Else data instance is self
Start: 5 Repeat the step 2 and 3 for all A_L
Start: 6 Compare the voting set V with A_L and find the DR and FAR
Output: Data instance having intrusion

5.1 Performance Metrics
DR, FAR and F1-Score are three metrics used to test the efficacy of the proposed technique. DR identifies the rightly classified anomaly by the system, FAR identifies the self is identified as non-self and F1-score measure the predictive power of any classification model. High DR and low FAR are the pre-requirement for any good anomaly detection technique.
1. Detection rate (DR): $TPTP+FN*$ 100, DR is defined as the total number of detected non-self when they are actually non-self.
2. False Alarm Rate (FAR): $FPFP+TN*100$, FAR is defined as the total number of detected non-self when they are actually self.
3. F1-Score: 2 $Precisio\ .RecallPrecision+Recall$, it is a measure of the harmonic mean of precision and recall which represents the predictive power of any classification model.

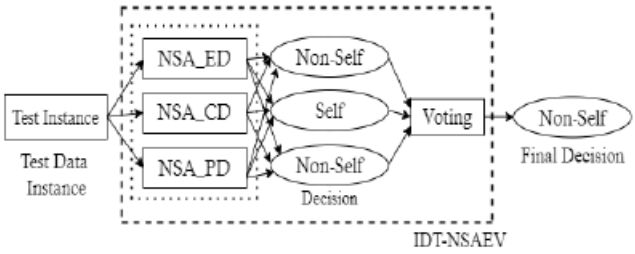

Figure 7: Flow diagram of ensemble voting algorithm.

Dimensionality Reduction
This work has used the column standardization technique for normalization of dataset, followed by proposed hybrid dimensionality reduction technique i.e., combination of stacked autoencoders feature extractor method and random forest feature selection method. Stacked autoencoders reduces from 41 features in NSL-KDD to 30 features. Next, the application of random forest feature selection method reduces the features further from 30 to 12. This reduction in dimensionality of dataset will lead to reduction in computational complexity for further processing.

Selection of Stable Threshold
Table 1 shows the variations in DR and FAR by changing with the Binding threshold values. As shown in table, the NSA_ED has 92.68% DR and 28.79% FAR at affinity value 0.35. The highest value of DR is 94.8% with minimum FAR (18.97%) at affinity value 0.4. Beyond this, with the increase of Binding threshold value, the DR decreases and FAR also increases. Similarly, for NSA_PD the average value of Binding threshold is 0.45 at which it gives highest DR (95.27%) with 23.56% FAR, and for NSA_CD the affinity value 0.55 gives highest DR (94.4%) with 28.07 FAR. From the results, binding threshold 0.40 for NSA_ED, 0.55 for NSA_CD and 0.45 for NSA_PD is chosen for further performance evaluations.

ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

Table 1: Optimum threshold value selection for NSA_ED, NSA_CD and NSA_PD.

Binding Threshold	NSA_ED		NSA_CD		NSA_PD	
	DR	FAR	DR	FAR	DR	FAR
0.35	92.68	28.79	80.23	43.23	93.22	54.3
0.40	**94.8**	**18.97**	87.44	26.96	89.36	24.64
0.45	90.19	24.94	90.71	26.44	**95.27**	**23.56**
0.50	86.86	23.19	91.09	27.19	88.63	23.56
0.55	80.55	30.26	**94.4**	**28.07**	83.91	24.71
0.60	73.29	40.67	94.78	59.4	47.36	7.16

5.2 Selection of Number of Detectors

Table 2 illustrates the results obtained by NSA_PD, NSA_CD and NSA_ED, by changing the number of detectors. The values in the table demonstrated that as the number of detectors is 20 in NSA_ED, the obtained DR is 94.8% and FAR is 18.97%, which is 156 The International Arab Journal of Information Technology, comparatively better performance as we increase the number of detectors. As the number of detectors increases the DR decreases and FAR rate increases. Similarly, NSA_CD gives the stable performance at 20 detectors, and NSA_PD has highest DR and lowest FAR at 25 number of detectors as shown in table. Beyond this, As the number of detectors is increased, the performance goes down. From the results, the number of detectors 25 for NSA_PD, and 20 for NSA_ED as well as for NSA_CD has been chosen for further performance evaluations.

Table 2: Selection of number of Detectors for NSA_PD, NSA_CD and NSA_ED.

Detectors	NSA_ED		NSA_CD		NSA_PD	
	DR	FAR	DR	FAR	DR	FAR
10	91.37	22.16	91.34	32.30	92.30	26.34
15	94.71	25.62	92.77	31.31	94.35	25.77
20	**94.8**	**18.97**	**94.41**	**28.07**	94.07	28.97
25	94.27	29.52	92.5	30.15	**95.27**	**23.56**
30	90.55	31.76	91.59	31.16	93.77	29.24
35	87.41	27.1	91.20	30.24	91.82	27.76
50	87.36	28.65	90.08	28.48	90.53	28.21

5.3 Performance of Proposed Technique (IDT-Nsaev)

Table 3 illustrates the performance of proposed technique in terms of DR, FAR and f1 score. Since, the proposed model combines the predictive power of three algorithms namely; NSA_ED, NSA_CD and NSA_PD. It increases the DR to 97.52% which is highest among the state of the arts techniques. Also, it reduces the FAR to 11.67% which is lower than NSA based algorithms compared in the table. Hence, it is experimentally evident that the proposed IDT-NSAEV technique enhances the prediction power by increasing the DR.

Table 3: Results of the proposed ensemble voting algorithm in comparison of the other three algorithms.

Algorithm	DR	FAR	f1 score
Proposed IDT-NSAEV	97.52%	11.67%	0.864865
NSA_PD	95.27%	23.56%	0.859544
NSA_ED	94.08%	18.97%	0.841142
NSA_CD	94.40%	28.07%	0.8202

ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

Table 4: Comparison of proposed Intrusion detection technique with other related techniques.

Technique	SM	Dataset	DR	FAR
Proposed IDT-NSAEV	ED, PD and CD	NSL-KDD	**97.52%**	**11.67%**
NSA-ED	ED	NSL-KDD	94.08%	18.97%
NSA-PD	PD	NSL-KDD	95.27%	23.56%
NSA-CD	CD	NSL-KDD	94.40%	28.07%
NSA with GA and DCN [2]	ED	----	81.70%	----
NADNS [7]	----	NSL-KDD kyoto2006+	96%	18%
AIS inspired IDS based on GA [3]	ED, MD	NSL-KDD	81.74 % -ED 77.44%-MD	----
RS based AIRS [19]	----	NSL-KDD	39.89%	----
MA-AIS [1]	----	NSL-KDD	89.78%	12.67%
MAIS-IDS [20]	----	NSL-KDD	90.54%	29.72%

VI. COMPARISON WITH THE RELATED WORK

To demonstrate the efficacy of the proposed IDT-NSAEV technique, a comprehensive comparison was conducted against other state-of-the-art intrusion detection techniques, specifically focusing on Detection Rate (DR) and False Alarm Rate (FAR). The results, summarized in Table 4, illustrate the performance of the proposed technique relative to existing NSA-based intrusion detection methods when applied to the NSL-KDD dataset. The data reveal that the proposed method achieved an impressive average DR of 97.52%, which is a notable improvement of 1.52% over the best-performing state-of-the-art techniques. Furthermore, the proposed technique significantly reduced the FAR to 11.67%, which is considerably lower than that achieved by the other methods. This substantial improvement in both detection accuracy and reduction of false alarms underscores the robustness and reliability of the IDT-NSAEV technique, highlighting its potential as an advanced solution for enhancing the security and effectiveness of intrusion detection systems in computer networks.

VII. CONCLUSIONS

This work proposes an innovative intrusion detection technique for computer networks, named IDT-NSAEV, which leverages an ensemble voting method using the Negative Selection Algorithm (NSA). The approach begins with a hybrid dimensionality reduction process that utilizes Stacked Autoencoders (SAEs) followed by a random forest algorithm to effectively reduce the dimensionality of the data. Following this, the NSA is applied using two key measures: Censoring Distance (CD) and Pattern Detection (PD), to evaluate their predictive capabilities in intrusion detection. The technique was tested on the NSL-KDD dataset and compared with the traditional NSA method based on Euclidean Distance (ED). The analysis of the results indicated that while the average performance of these algorithms in terms of Detection Rate (DR) was comparable, they exhibited inconsistent behavior individually. To address this issue, the proposed IDT-NSAEV technique combines the predictive strengths of the NSA_ED, NSA_CD, and NSA_PD algorithms, resulting in a more robust and stable intrusion detection system. The proposed method achieved an impressive average Detection Rate of 97.52%, outperforming state-of-the-art techniques using NSA by 1.52%. Additionally, the proposed technique demonstrated a significant reduction in the False Alarm Rate (FAR). For future work, there is potential to enhance this approach further by incorporating other similarity measures, such as Manhattan Distance and Minkowski Distance, with the proposed ensemble technique, thereby potentially improving its efficacy and stability in intrusion detection applications.

ISSN(Online): 2320-9801

ISSN (Print): 2320-9798

International Journal of Innovative Research in Computerand Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Vol. 4, Issue 1, January 2016

REFERENCES

1. Aziz A., Hanafi S., and Hassanien A., "Multi-Agent Artificial Immune System for Network Intrusion Detection and Classification," in Proceedings of International Joint Conference SOCO'14-CISIS'14-ICEUTE'14, Bilbao, pp. 145-154, 2014.
2. Aziz A., Salama M., Hassanien A., and Hanafi S., "Detectors Generation Using Genetic Algorithm for A Negative Selection Inspired Anomaly Network Intrusion Detection System," in Proceedings of Federated Conference on Ensemble Voting based Intrusion Detection Technique using Negative Selection Algorithm 157 Computer Science and Information Systems, Wroclaw, pp. 597-602, 2012.
3. Aziz A., Salama M., Hassanien A., and Hanafi S., "Artificial Immune System Inspired Intrusion Detection System using Genetic Algorithm," Informatica, vol. 36, no. 4, pp. 347-358, 2012.
4. Asuvaran & S. Senthilkumar, "Low delay error correction codes to correct stuck-at defects and soft errors", 2014 International Conference on Advances in Engineering and Technology (ICAET), 02-03 May 2014. doi:10.1109/icaet.2014.7105257.
5. Balachandran S., Dasgupta D., Nino F., and Garrett D., "A Framework for Evolving Multi-Shaped Detectors in Negative Selection," in Proceedings of IEEE Symposium on Foundations of Computational Intelligence, Honolulu, pp. 401-408, 2007.
6. Forrest S., Perelson A., Allen L., and Cherukuri R., "Self-Nonself Discrimination in A Computer," in Proceedings of IEEE Computer Society Symposium on Research In Security and Privacy, Oakland, pp. 202-212, 1994.
7. González F., Dasgupta D., and Gómez J., "The Effect of Binary Matching Rules in Negative Selection," in Proceedings of Genetic and Evolutionary Computation Conference, Chicago, pp. 195-206, 2003.
8. Gonzalez F., Dasgupta D., and Niño L., "A Randomized Real-Valued Negative Selection Algorithm," in Proceedings of International Conference on Artificial Immune Systems, pp. 261-272, 2003.
9. Ji Z. and Dasgupta D., "Real-Valued Negative Selection Algorithm with Variable-Sized Detectors," in Proceedings of Genetic and Evolutionary Computation Conference, USA, pp. 287-298, 2004.
10. Ji Z. and Dasgupta D., "Applicability Issues of The Real-Valued Negative Selection Algorithms," in Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, Seattle, pp. 111-118, 2006.
11. Ji Z. and Dasgupta D., "Revisiting Negative Selection Algorithms," Evolutionary Computation, vol. 15, no. 2, pp. 223-251, 2007.
12. Sanobar, K. and Vanita, M. (2013) SQL Support over MongoDB Using Metadata. International Journal of Scientific and Research Publications, 3, 1-5. https://technet.microsoft.com/enus/library/bb522562(v=sql.105).aspx
13. McHugh J., "Testing Intrusion Detection Systems: A Critique of The 1998 and 1999 Darpa Intrusion Detection System Evaluations as Performed by Lincoln Laboratory," ACM Transactions on Information and System Security, vol. 3, no. 4, pp. 262-294, 2000.
14. Wei-Ping, Z. and Ming-Xin, L. (2011) Using MongoDB to Implement Textbook Management System instead of MySQL. 3rd IEEE International Conference on Communication Software and Networks (ICCSN), Xi'an, China, 27-29 May 2011. https://doi.org/10.1109/ICCSN.2011.6013720.
15. Powers S. and He J., "A Hybrid Artificial Immune System and Self-Organising Map for Network Intrusion Detection," Information Sciences, vol. 178, no. 15, pp. 3024-3042, 2008.
16. Poggiolini M. and Engelbrecht A., "Application of the Feature-Detection Rule to the Negative Selection Algorithm," Expert Systems with Applications, vol. 40, no. 8, pp. 3001-3014, 2013.
17. Sabri F., Norwawi., and Seman K., "Hybrid of Rough Set Theory and Artificial Immune Recognition System as A Solution to Decrease False Alarm Rate in Intrusion Detection System," in Proceedings of 7th International Conference on Information Assurance and Security, Melacca, pp. 134-138, 2011.
18. Seresht N. and Azmi R., "MAIS-IDS: A Distributed Intrusion Detection System Using Multi-Agent AIS Approach," Engineering Applications of Artificial Intelligence, vol. 35, pp. 86-298, 2014.