Intrusion Detection System for IoE-Based Medical Networks

Parul Lakhotia, Netaji Subhas University of Technology, Delhi, India

Rinky Dwivedi, Maharaja Surajmal Institute of Technology, Delhi, India

Deepak Kumar Sharma, Indira Gandhi Delhi Technical University for Women, Delhi, India*

Nonita Sharma, Indira Gandhi Delhi Technical University for Women, Delhi, India

ABSTRACT

Internet of everything (IoE) has the power of reforming the healthcare sector - various medical devices, hardware, and software applications that are interconnected, tendering a massive volume of data. The huge interconnected medical-based network is prone to significant malicious attacks that can modify the medical data being communicated and transferred. IoE permits dynamic two-way communication and empowers the network with intellect, sophisticated data handling, caching, and allocation mechanisms. In this paper, an improvement in the conventional variable-sized detector generation for healthcare - IVD-IMT algorithm under Artificial Immune System (AIS) based Intrusion Detection System (IDS) capable of handling enormous data generated by the IoE medical network is proposed. Algorithm efficiency is dependent on two performance metrics - detection rate and false alarm rate. The input parameters were tuned using synthetic datasets and then tested over the NSL-KDD dataset. The research lays emphasis on lowering the false alarm rate without compromising on the detection rate.

KEYWORDS

Internet of Everything, Medical Networks Security, Intrusion Detection System, Real Time, Artificial Immune System, Multidimensional Data Points

1. INTRODUCTION

The Internet of Everything (IoE) marks a step forward from the Internet of Things (IoT). IoT connects various devices into internet-like networks such as RFID or NFC, to enable end-to-end data transfer from users to the cloud. It is one way of communication with the sole purpose of collecting data from the environments these gadgets are placed in. On the other hand, IoE supports bidirectional communication through intelligent networks which include both gadgets and people. The bottom layer is aware of its environment and does more than mere data collection.

With advancements in IoE, millions of devices over the internet can conduct communication. IoE (Miraz et al., 2018; Patel & Patel, 2016; Ryan & Watson, 2017) has its application in the healthcare

DOI: 10.4018/JDM.321465

*Corresponding Author

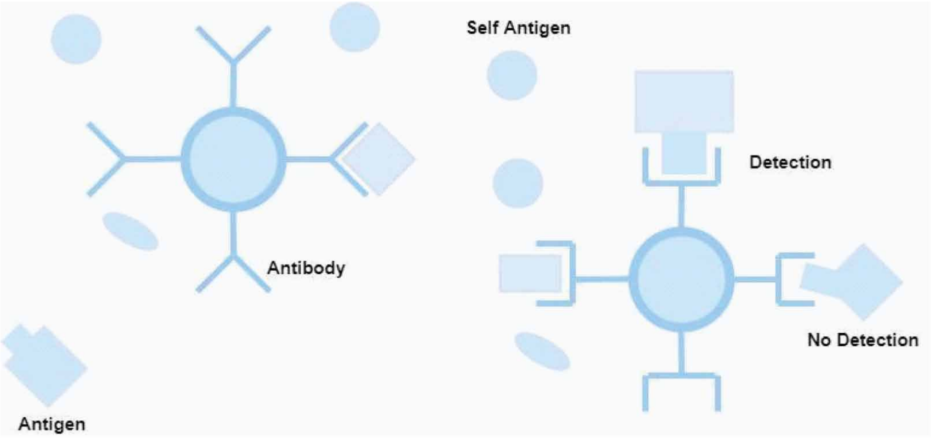
sector, providing real-time services with reduced healthcare costs. It has enhanced the performance, precision, and accuracy of medical procedures.

Medical networks (Dong et al., 2016; Dwivedi et al., 2019; Kumar & Bairavi, 2016) comprise Internet of Medical Things (IoMT) devices that record patient information, actuators that display results, processing units to generate reports, and finally data management units such as cloud storage. Such complex systems that include manual operations and automation at higher levels see high traffic influx on a day-to-day basis. Network security in such cases poses a fair challenge for two reasons - latency and accuracy. While any security system needs to be robust, a breach of medical data may cost a patient their life, data rate cannot be compromised in serious health situations. Security checks are required to be highly optimized to deliver real-time data. If the security system has high sensitivity, any legitimate change in normal state may raise numerous false alarms. The healthcare sector has faced a massive number of cybersecurity attacks in recent decades. The security of a network majorly focuses on authentication, confidentiality, and integrity (Muhammad et al., 2017; Yeole & Kalbande, 2016). Existing Intrusion Detection Systems (IDS) (Foley, 2021; Sunke, 2008; Tiwari et al., 2017; Xu et al., 2013) are prone to stealthy attacks like Man in the Middle attack, where parameters like CPU usage and loop latency see a negligible change and end up undetected. Dynamic networks like IoE require dynamic security systems that can adapt to the new normal seamlessly, without compromising on the detection rate.

Various methods have been proposed for the implementation of IDS, the Artificial Immune System (AIS) (Balthrop et al., 2002; Dasgupta et al., 2004; Read et al., 2012) being one of them. Biological immune systems have antibody cells called lymphocytes that provide immunity to the body from pathogens. These antibody cells are closely modelled as detectors in the AIS and have the same properties as lymphocytes and other antibodies. AIS integrates the principles and fundamentals of the biological immune system (Srivastava & Lin, 2021), incorporating error resistance, dynamic adaptation, real-time self-detection, and computational facilities. Lymphocytes are referred to as negative detectors as they are qualified for binding to non-self-cells.

Like all predictive models, AIS can produce false results in the form of false negatives and false positives. A high false positives value would indicate autoimmunity, while a high false negatives count brings the detection rate down. This paper lays emphasis on optimizing the generation of dynamic detectors, using Negative Selection; that can distinguish between non-self and self-cells. Figure 1 depicts a simplified diagrammatic version of the artificial immune system, where specific detectors are generated to only detect non-self-antigens. This paper attempts to generate detectors in

Figure 1. Artificial Immune System generated by multiple dimensions in terms of detection rate and false alarm rate



a multidimensional space, with each dimension representing a parameter that categorizes any point in space into self and non-self.

The main contributions to this paper are as follows:

1. Improve the time complexity of the Variable sized detector generation algorithm during detector generation through the proposed IVD-IMT algorithm to handle the enormous data generated by medical networks established by the IoE model.
2. To ensure the security of voluminous and sophisticated data, better coverage around self or non-anomalous data points to better detect stealthy attacks that may pass as normal in physical anomaly detection systems.
3. Draw a comparative study of variable-sized detector generation for the massive statistics

The structure of the paper is as follows; Section 2 enlists the previous work on security in medical networks. Section 3 discusses our motivation for the paper. Section 4 discusses the proposed AIS approach for implementing IDS. Section 5 discusses the results followed by section 6 which concludes the study.

2. RELATED WORK

AIS has been known to implement Intrusion detection systems for ensuring a secure network in the field of healthcare. There have been several models proposed to address security issues. Yang et al. (2014) have proposed a detailed description of implementing an IDS system based on the AIS. It comprises methods like encoding antibodies, evolution mode, and generation algorithms. A detailed summary regarding the IDS of IoT implemented on the Negative selection algorithm (NSA) and Danger theory was proposed by Pamukov et al. (2017) along with a comparative analysis on the same. The paper also outlined the prerequisites required for the IoT IDS. An amalgamation of AIS and genetic algorithms had been proposed by Barani et al. (2014) detecting instructions that are dynamic in nature and present in mobile networks.

David J. Langley et al. (2021) explore the fundamentals of IoE models, and their value addition to running businesses. They talk about the levels of smartness that the components of a network may be endowed with and their effect on the overall network communication. The devices can now do more than just send the data they receive, turning the communication into bidirectional activity.

Ying Tan et al. (2016) discuss the AIS and human immune system, it further recognizes and studies the concepts related to the computer immune system (CIS). There were several papers presented that were based on anomaly detection. Angelov et al. (2016) discussed the state-of-art of AIS and reviewed various immune established algorithms and functions, further discussed the associated application, and contributed to the establishment of vigorous IS-based algorithms. It also outlines the various AIS approaches that exist. R. Banu et al. (2016) presented a paper that introduced various methods for ensuring security in IoT based on the biological immune system. It establishes that biologically based models for ensuring security provide vigorous defense and decentralized systems. These models provide a more scalable system and are self-organized. Mahdi H. Miraz et al.(2021) talk about the expanding use cases of IoE in corporations and the impact they can have on customer experience in current ecosystems. Bayar et al. (2015) propose biological immune system-inspired methods for fault detection, diagnosis, and recovery (FDDR). It summarizes the essentials for FDDR and highlights biological immune systems and approaches with regard to FDDR issues. It distinguishes the AIS into three divisions such as one-signal, immune network, and two-signal-inspired methods. A probability investigation was carried out to find the association between the number of detectors and the identification probability of a random fault. This was studied by D'haeseleer et al. (1996) and utilized matching probability for assessing the number of detectors. This paper discusses the approaches that are statistical in nature for analyzing the coverage of the detector in NSA, known as an evaluation that is quantitative. V-detectors are known to eliminate the detector coverage issues with effective methods by Z. Ji et al. (2005). V-detectors handle the issue of detector coverage differently with innovative techniques by estimating the area covered by the definite set of detectors. Dipankar et al. (2004) proposed V- detector NSA for addressing the detector coverage problem. The solution focuses on calculating proximate coverage on the generation of detector sets. It also highlights that when detectors of constant size are generated, the detector count needs to be mentioned beforehand.

3. MOTIVATION

IoE networks have tremendous potential in the medical field. However, in practice, these networks need strong protection against intrusion. They transfer highly sensitive medical information. With increased applications of the medical network in healthcare, the possibilities of malicious attacks on the network have increased. Conventional methods of placing physical intrusion detection nodes in the network make the orchestration less dynamic and pass stealthy attacks as normal network behavior. This served as the motivation to design a secure IoE-based intelligent intrusion detection system for the medical network with minimal overhead and a high detection rate. Multiple false alarms from a highly sensitive algorithm may lead to shutting down the entire alerting system, defeating the very purpose of subtle anomaly detection. AIS presents itself as an innovative solution to maintain this balance. It trains on the set of ‘self’ data points provided by the user, to prevent autoimmunity.

4. PROPOSED MODEL

This section contains a detailed analysis of the IVD-IMT algorithm, in terms of time complexity, and detector generation phases and maps them to the AIS principles used. It also talks about the dataset used along with feature extraction implementation to improve model performance.

4.1 Negative Selection Algorithm With Detectors of Variable Size

IVD-IMT is based on the Negative Selection Algorithm (NSA) (Bendiab & Kholladi, 2010; Igawa & Ohashi, 2009; Yang et al., 2020), an integral part of AIS. It distinguishes between self and non-self-cells. NSA comprises a detector set that is generated during the training phase using the self-cells and autoimmune detectors are removed. An advantage of using this technique over neural networks comes from the adaptive nature of the generation process. If new types of cells are identified as self in the future, generators surrounding that area can be regenerated to accommodate the changes instead of training the entire model again. Fundamental goals for all negative selection algorithms are: (1) reducing the number of detectors generated, to maintain a lightweight model; (2) ensuring that the set of detectors addresses maximum possible anomalies, in other words, to maximize no self-space coverage; (3) efficient generation of detector sets, in terms of time complexity and computational power required by the generation algorithm. IVD-IMT has been built upon the idea of detectors with variable sizes. Detector generation is considered complete when target coverage, an input parameter to the algorithm, is reached. Variable-sized detectors (Ji & Dasgupta, 2004; Ji & Dasgupta, 2009; Beyer, Shapiro, Lamont et al, 2005) have an associated advantage in which the detectors’ geometry is of no relevance. Hence, it does not lead to any difficulties in using various representations for detectors.

In most real-valued NSA algorithms, the detector size is decided according to a predefined threshold. Variable-sized detectors do not work on such constraints, which optimizes the generation algorithm. The radius of each detector depends on the position of the nearest self-point in the space. In areas sparsely populated by self-points, detector size grows to ensure maximum coverage per detector. For estimating the area covered, the algorithm utilizes the point percentage enclosed by the detectors.

4.2 Dataset Description

The data set used for training is the NSL-KDD (Meena & Choudhary, 2017) training data set. It comprises the records selected from the complete KDD data set. The training data set does not contain any redundant or duplicate data points. There are 21 attacks present in the training data-set and there exist 37 attacks in the testing data set. The different attacks that are known are contained by the training data set whereas the standard attacks are the extra attacks that the test data set consists of. The attacks have been classified into four major classes: dos, probes, privileges, and access attacks. The below section lays emphasis on the feature extraction method. Figure 2 depicts the distribution of the NSL-KDD data set.

4.3 Feature Extraction - Principal Components Analysis (PCA)

Principal component analysis (Hidayat et al., 2011; Lhazmir et al., 2017; Murali, 2015) is a feature extraction method utilized for decreasing the features and dimensions for improving the computational ability of the model. The data set usually comprises several associated variables, signifying possibilities of various redundant variations. PCA ensures the retention of relevant features in the data set by converting to a group of new variables which are linear functions of the original data referred to as principal components (PCs), these are not correlated and are oriented in a specific order such that only the initial contains significant features composed by the original features. Deciding on the new variables boils down the problem to only determining the eigenvalue/vector; this makes PCA an adaptive data analysis technique. This method is successful in maximizing the variance and retaining the variance to a great extent. The section below describes the algorithm. Figure 3 (a): Distribution of NSL KDD data-set for 4 PCA components according to attack flag - anomalous and non-anomalous data points and figure 3 (b): Distribution of NSL KDD data-set for 4 PCA components according to attack map - 4 types of attacks - dos attacks, probe attacks, privilege attacks, access attacks.

Figure 2. Distribution of NSL KDD data-set

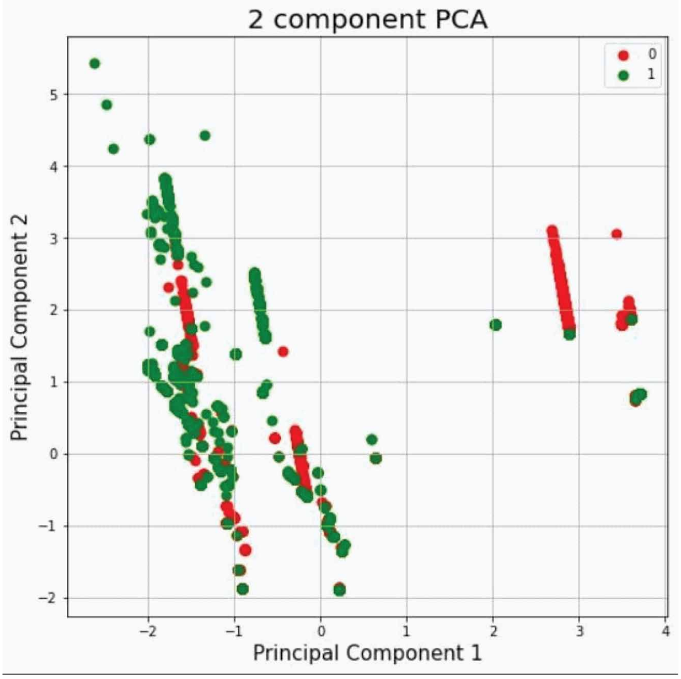
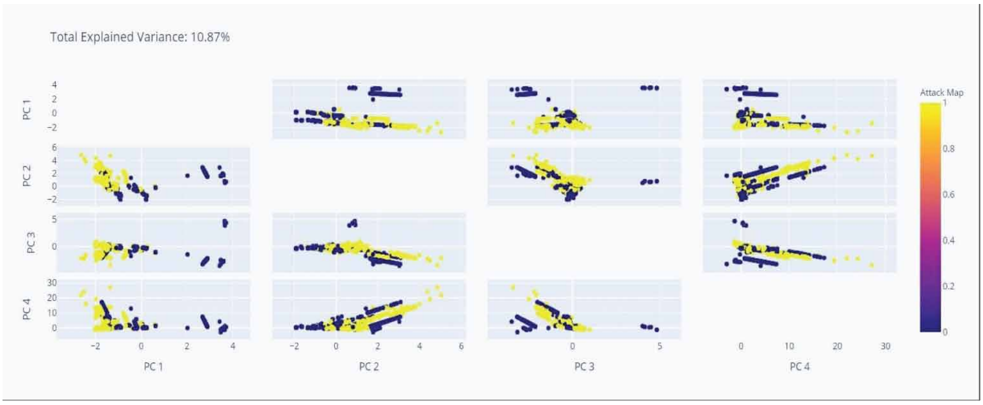


Figure 3. Distribution of NSL KDD data-set for 4 PCA components according to attack flag - anomalous and non-anomalous data points



4.4 Proposed Algorithm for Detector Generation

This section talks about the acceptance criteria, i.e., when the algorithm has generated enough detectors. It also proves the optimization in terms of time complexity over the traditional V-Detector algorithm, followed by the IVD-IMT algorithm.

4.4.1 Coverage

For a given detector set in a defined region, coverage can be calculated as the ratio of the area of the non-self region (Ji & Dasgupta, 2004; Beyer, Ji, & Dasgupta, 2005) covered by detectors to the total area of the non-self region. If the region is taken as a set of points, it can alternatively be seen as the number of unsuccessful attempts to find an uncovered point in the non-self region to the total number of attempts. Therefore, if the algorithm finds an uncovered point at the n^{th} attempt, by picking one random point at a time, the total number of unsuccessful attempts becomes n-1. Equation 1 provides the total detector coverage. Hence:

$$C = \frac{n-1}{n} \tag{1}$$

where C is the total detector coverage in the non-self region. It can be given as a parameter to the algorithm in the form of the minimum coverage required.

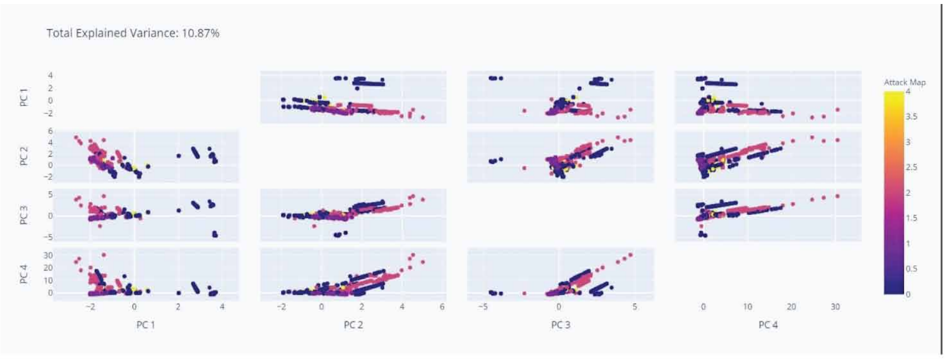
The algorithm generates enough detectors such that the coverage is equal to or greater than C.

4.4.2 Detector Generation

The idea is to divide the entire non-self region into groups, to improve the coverage achieved by v detectors (or variable-sized detectors) while reducing the time complexity by saving iterations through the entire self-set for every single detector. The areas near self-points need good coverage to detect stealthy attacks that may otherwise go unnoticed due to minimal changes in network parameters like loop delay or additional computational power (Forrest et al., 1994; Hart, 2005). Figure 4 displays the detector generation in GUI. This algorithm can be divided into two major steps:

1. To identify self-region clusters or groups in the entire space, regions with concentrated densities of self-set points and group them into self-regions, each with a centre as the mean of coordinates of the constituent points. Once all central points have been identified, the detector generation

Figure 4. Distribution of NSL KDD data-set for 4 PCA components according to attack map - 4 types of attacks - dos attacks, probe attacks, privilege attacks, access attacks



is initiated for every group centre. Detectors are generated in the vicinity sphere, i.e. within a specified radius from the respective centre, to closely cover the boundaries of these groups. The radius should cover the farthest self-point in the group, as measured from the centre of that group. The degree of closeness to the self-set boundary can also be supplied as a parameter. If the self-threshold is too small, the space between self-samples could not be represented. In other words, more samples are needed to train the system properly. On the other hand, if the self-threshold is large, the false self region represented by the boundary samples may be too large to accept.

2. After iterating through every group centre, random points in the non-self region are chosen as detector sphere centres, and their radius is taken as the centre of the nearest detector found in the current detector set. This would result in a huge improvement over finding the nearest self-set point for every detector in terms of time complexity for large datasets of self-data since the size of the detector set would be of a much smaller order. The time complexity of the detection generation algorithm can be calculated as follows by equations 2 and 3:

$$O(n) = \sum_{i=1}^{r}\left(d_i \times S_i\right) + D_i \times \left(\sum_{i=1}^{r} d_i\right) \tag{2}$$

$$D = D_i + \sum_{i=1}^{r} d_i \tag{3}$$

where r is the number of groups, S_i is the number of self points in that region, d_i is the number of detectors in the vicinity sphere and D_i denotes the number of detectors generated after the first phase of the algorithm is over, that is detectors lying in the region outside the vicinity sphere, D is the total size of the detector set.

This time complexity can be proved to be less than that of the conventional algorithm for detectors of variable size. The algorithm has been built on the assumption that the size of the detector set will always be less than the total number of self-points in the dataset. For the conventional algorithm, every detector is generated with a radius equal to the nearest self point, which would need an iteration over the entire self-set in the worst case, hence for D number of detectors, it can be represented as:

$$O(n) = D \times \sum_{i=1}^{r} S_i$$

$$= \sum_{i=1}^{r} d_i \times \sum_{i=1}^{r} S_i + D_i \times \left(\sum_{i=1}^{r} S_i\right) \tag{4}$$

Now:

$$\sum\left(A \times B\right)\left\langle \sum A \times \sum B \ for \ A\right\rangle 1 \text{ and } B > 1$$

4.4.3 Design Architecture for IVD-IMT

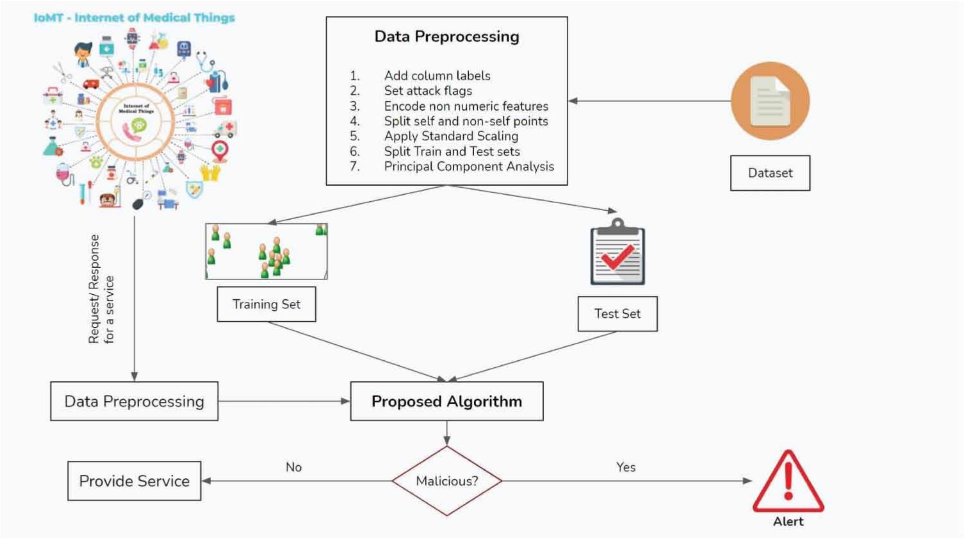
In the following analysis, it was assumed that both self and non-self points appear in some bounded n-dimensional real space. Some finite numbers of self-samples are provided as input (Ji & Dasgupta, 2009). They are randomly distributed over the self-region. The training data is noise-free, meaning all the self-samples are real self-points. This is not necessary for principle but is used to simplify the discussion. To evaluate the detection performance, the testing data are a finite number of random points over the entire space in the question described above. Each of those points can be verified to be self or non-self. Figure 5 below depicts the proposed architecture for IoE environment. The process begins with capturing data from the patient, pre-processing to make it fit for feeding to the detector generation algorithm, and finally testing the model to report intrusion.

4.4.4 The IVD-IMT Algorithm

Radius of the vicinity sphere is calculated by the given formula in equation 5 as Euclidean distance between the mean of coordinates in the group and the farthest self point:

$$Radius \ of \ vicinity \ sphere = \sqrt{\sum_{i}^{n}(u_i - l_i)^2} \tag{5}$$

Figure 5. Architecture of IVD-IMT for IoE network environment


IoMT - Internet of Medical Things
Internet of Medical Things
Data Preprocessing
1. Add column labels
2. Set attack flags
3. Encode non numeric features
4. Split self and non-self points
5. Apply Standard Scaling
6. Split Train and Test sets
7. Principal Component Analysis
Dataset
Training Set
Test Set
Request/ Response for a service
Data Preprocessing
Proposed Algorithm
Malicious?
No
Yes
Provide Service
Alert

Algorithm 1. IVD-IMT

```
   Data: target coverage, threshold, self set
   Result: Detector set for the supplied self set$y = x^n$
/* Phase 1 - detector generation inside vicinity spheres
   Divide the self set into groups or clusters of closely spaced self points currentdetectorset ¬[ ];
   for each group do local number of covered points ¬ 0 ;
      Use equation (4) to calculate the radius of vicinity sphere; while True do
         Select a random point P in the vicinity sphere as the centre of the detector; if P is a self point then continue
         end
          else
            if P is non self and uncovered then
            /* Create new detector D
               D.Centre ¬Coordinates of P;
               D.Radius ¬Euclidean distance between P and nearest self point in vicinity sphere;
               current detector set¬ D;
            end
            if point is non self and covered then
           /* point is already covered
               number of local covered points += 1; if current coverage $^3$ targetcoverage then break;
               end
             end
          end
      end
   end
   /* Phase 2 - detector generation outside vicinity spheres
   Select a random point in the entire region ; if point is self point then continue;
   end
   if point is non self and point is uncovered then
       X ¬nearest detector in the current detector set. ;
      Radius ¬Euclidean distance between the point and X ;
      Add the generated detector to the current detector set; else
      number of covered points += 1
   end
if current coverage $^3$ targetcoverage then return current detector set end
```

where l_i is the coordinate of the farthest point in ith dimension and u_i is the coordinate of the mean or the group center in ith dimension and n is the number of dimensions.

The self-space in 2 dimensions (figure 6), as obtained after performing PCA to generate 2 components of the training data-set.

Axes have been inverted in the following GUI implementation, but the same data set has been used to generate circular variable-sized detectors (figure 7), for low and high target coverage respectively:

5. RESULTS AND ANALYSIS

The effect of the parameters of control and the variations in the strategy has been studied by means of additional experiments. The difference in results was found to be related to the number of sample points or a variety of forms in terms of self-region, including the specific geometric parameters. On the part of the algorithm, the difference was dependent on the target coverage and provided a threshold as the minimum distance from self-points. The following sections lay emphasis on the evaluation criteria and further analysis of the algorithm.

5.1 Evaluation Metrics

The NSL KDD data set was divided into training and testing sets, the training set consisting of self-points and the test set as a shuffled set of self and non-self points, in a 20:80 ratio respectively.

Figure 6. Diagrammatic representation of self-points for NSL-KDD

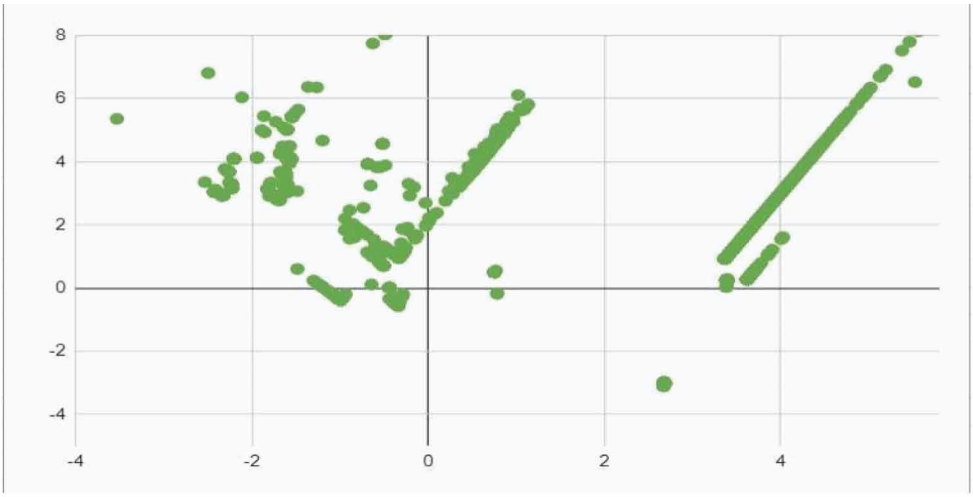
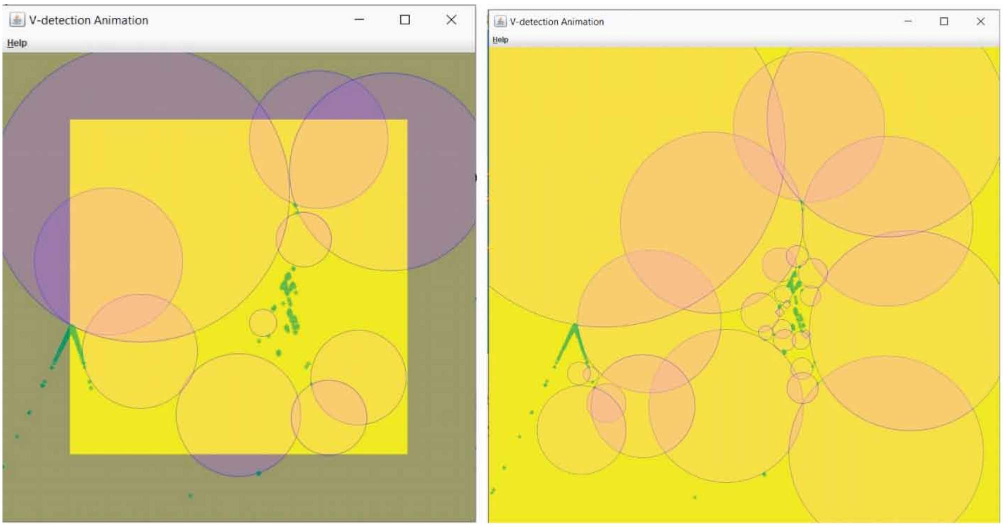


Figure 7. Detector Generation in GUI



DR(Detection Rate) and FAR(False Alarm Rate) are the criteria opted for evaluation. They are defined as:

$$\text{DR= nonself correctly classified/ total nonself} = \frac{TP}{TP + FN} \tag{6}$$

$$FAR = self\ \ incorrectly\ \ classified\ /\ totalself = \frac{FP}{FP + TN} \tag{7}$$

where TP stands for a number of True Positives, which means the test point belonged to the non-self region and was detected as an anomaly by the algorithm, FP stands for the number of False Positives, which means the test point belonged to the self region and was detected as an anomaly by the algorithm, FN stands for the number of False Negatives, which means the test point belonged to the non-self region and was not detected as an anomaly by the algorithm.

5.2 Graphical Analysis

The two-performance metrics, as discussed at the beginning of this paper - detection rate and false alarm rate were plotted against input parameters to IVD-IMT; namely the number of dimensions, target coverage, and the value of the threshold, which is a measure of how tightly bound the detectors must be to the boundaries of the self region. The results have been shown in Figure 8. It was observed that an increase in the dimension results in a fall in the detection rate and an increase in the false alarm rate as depicted in the following graphs. If the threshold is too high it results in decreased detection rate and defeats the purpose of the algorithm to detect stealthy attacks, while a value too low increases the false positive rate and causes unnecessary alerts. Figure 8 below depicts the graphical analysis.

Figure 8. Graphical Analysis

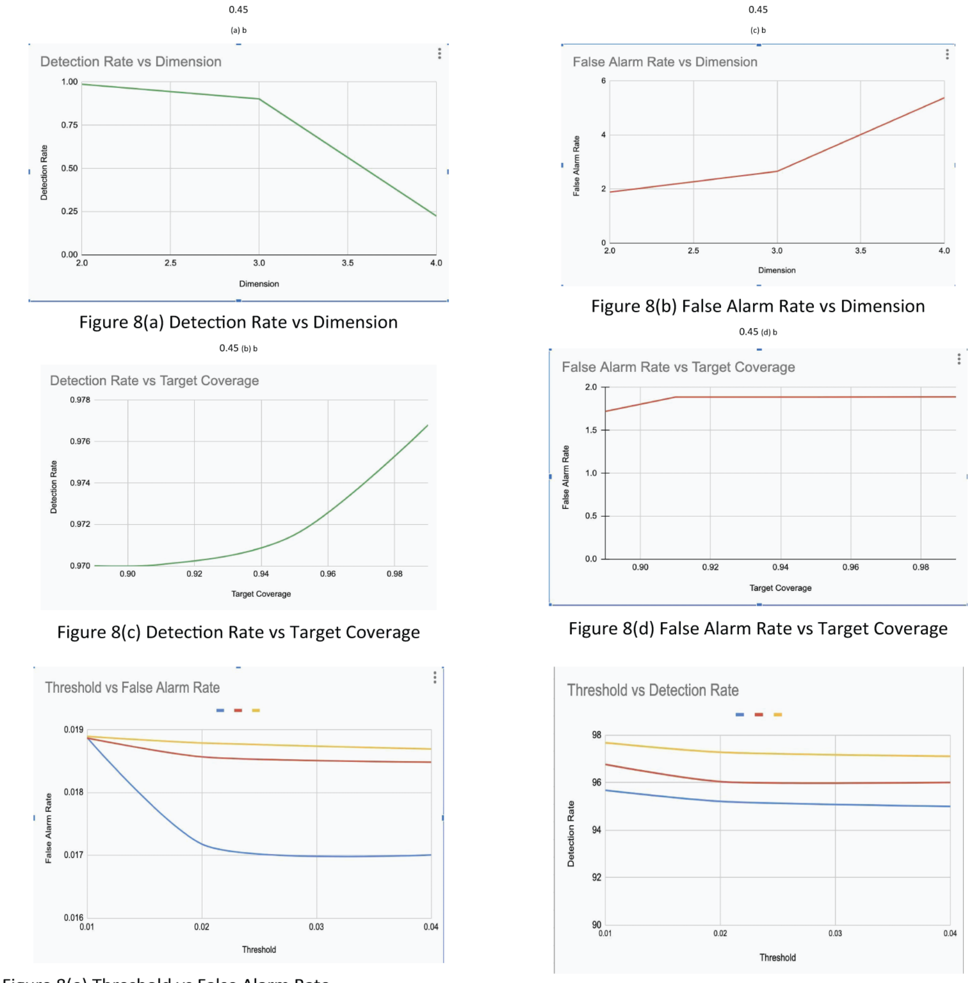


Figure 8(a) Detection Rate vs Dimension

Figure 8(b) False Alarm Rate vs Dimension

Figure 8(c) Detection Rate vs Target Coverage

Figure 8(d) False Alarm Rate vs Target Coverage

Figure 8(e) Threshold vs False Alarm Rate

Figure 8(f) Threshold vs Detection Rate

Starting from a detection rate of 97.767% and a false alarm rate of 1.88% for 99% target coverage and 0.01 threshold in 2-dimensional space, the deviation in detection rate at 0.56% decline was significantly less than the 36.7% rise in false alarm rate. A steep curve was observed in the 4-dimensional stage, with a relative decrease of 76.95% and 106.22% in detection rate and false alarm rate respectively. Therefore, the application of IVD-IMT should be restricted to three-dimensional data points. Varying the target coverage parameter showed minimal deviation in false alarm rate with an increasing curve since the higher the target coverage, the larger the area covered by detectors. In case the self-points in test sets do not fall under any vicinity spheres created around training self points data, these outliers may increase the false alarm rate. The detection rate shows a significant improvement as it is directly proportional to the target coverage value provided. The graphs were plotted for 2-dimensional data analysis, at a 0.01 threshold. Similar observations were made in the case of the threshold parameter, a high threshold indicates loose boundaries and hence the false alarm rate goes down. The detectors would be at a distance from the self-points and hence not detect any closely outlying self-points near the boundaries. The graphs were plotted for 99%, 95%, and 89% threshold values in yellow, red, and blue color respectively.

5.3 Comparative Analysis

Axes have been inverted in the following GUI implementation. Still, the same data-set has been used to generate spherical variable-sized detectors, to draw a comparative study of the conventional V detector algorithm vs the proposed version of the variable detector algorithm.

The self space in 2 dimensions, as obtained after performing PCA to generate 2 components of the training data-set. Figure 9 displays the representation of the self points in the self region. Figure 10 represents the detector generation in GUI for the V-detector and IVD-IMT algorithm.

Figure 10 shows the detector generation in the traditional V-Detector algorithm and IVD-IMT. The generation of detectors around the boundary can be easily distinguished in both cases, due to the formation of vicinity spheres, boundaries are densely populated with smaller-sized detectors which have an impact in increasing the detection rate. This would be especially advantageous in detecting stealthy attacks since the anomalous points would lie in proximity with the self region and would be passed as non-anomalous by physical intrusion detection models.

Figure 9. Diagrammatic representation of Self points in the test set

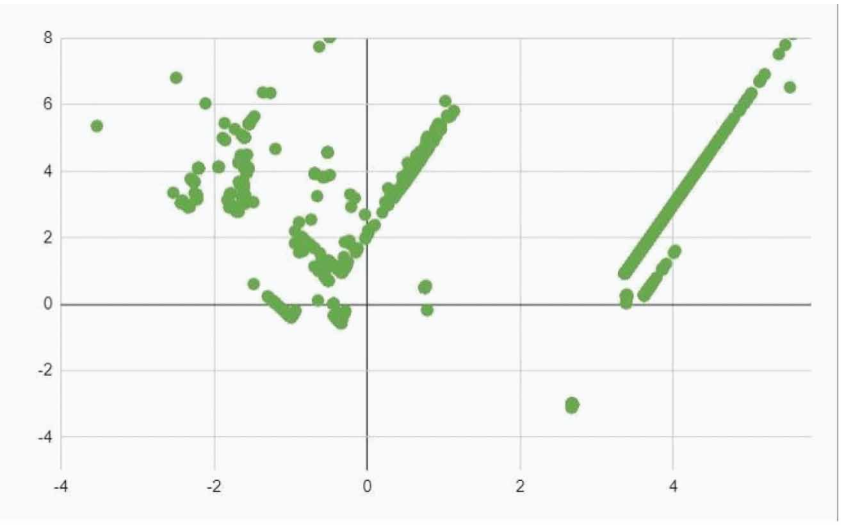


Figure 10. Detector generation in GUI for V detector and proposed algorithm

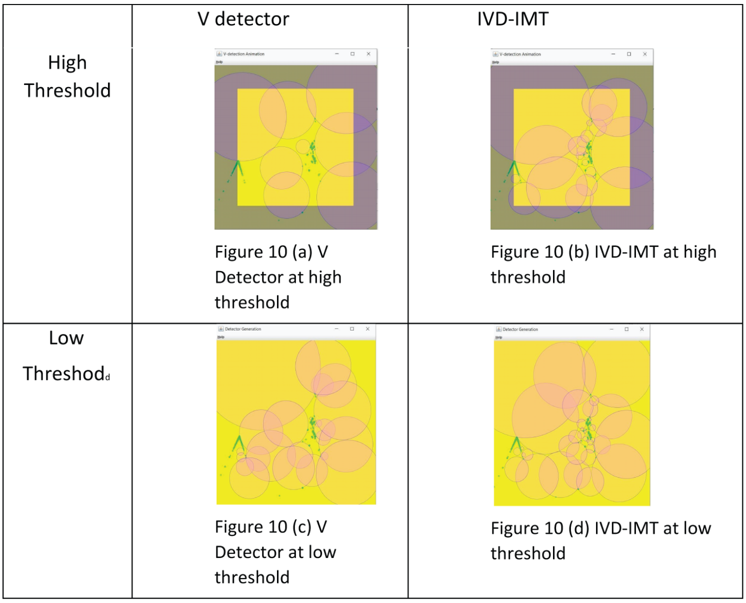


Figure 10 (a) V Detector at high threshold

Figure 10 (b) IVD-IMT at high threshold

Figure 10 (c) V Detector at low threshold

Figure 10 (d) IVD-IMT at low threshold

Table 1 compares the detection rate and false alarm rate accuracy of the IVD-MT algorithm with the conventional V Detector algorithm. The algorithm achieved coverage of 99%, with a threshold value of 0.01, i.e. closely covering the boundary of the self spaces with this limit set at 1000 detectors. Overall, a detection rate of 97.767% and a false alarm rate of 1.88% were achieved for 99% target coverage in 2-dimensional space which proves to be highly successful than the conventional V Detector. Table 2 compares the accuracy of IVD-MT with other Machine Learning algorithms (Thirumalai & Mohan, 2020). ML algorithms are supposed to be computationally heavy and are known to have a greater time complexity. IVD-MT is at par with other ML algorithms and in addition, provides better computation power with reduced time complexity. Figure 11 (a) represents the comparative Analysis of Detection rate and False Positive Rate and figure 11 (b) represents the comparative Analysis of time complexity.

Table 1. Comparative analysis with V detector and ML models

Name	Detection Rate	False Alarm Rate	Remarks
Linear Support Vector Machine Classifier	0.971	0.0386	Computationally heavy, greater time complexity
Quadratic Support Vector Machine Classifier	0.982	0.0768	Computationally heavy, greater time complexity
K-nearest-neighbor Classifier	0.953	0.0292	Higher False Alarm Rate
Linear Discriminant Analysis Classifier	0.939	0.0365	Higher False Alarm Rate
IVD-IMT (for 2-dimensional data points)	0.976	0.0188	Low time complexity, lowest false alarm rate, high accuracy
Conventional V Detector	0.782	0.026	Low Detection Rate, high false alarm rate, low accuracy

Figure 11a. Comparative Analysis of detection rate and false-positive rate with alternative models to detect intrusion

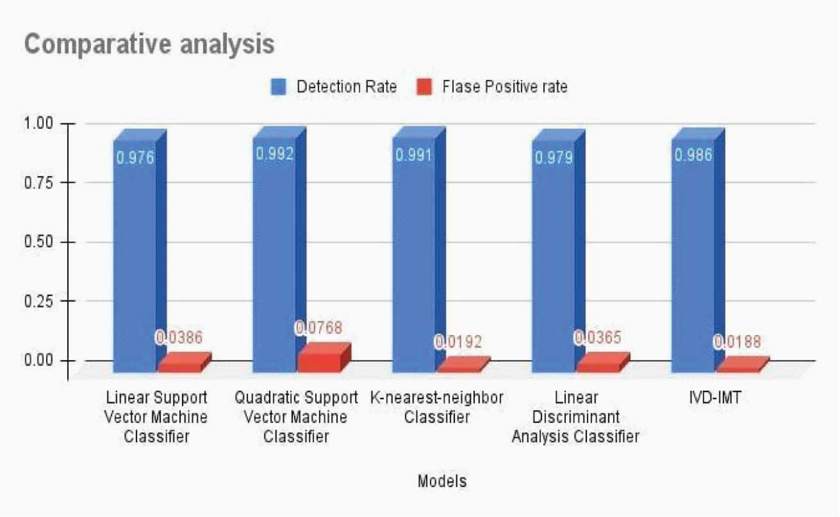
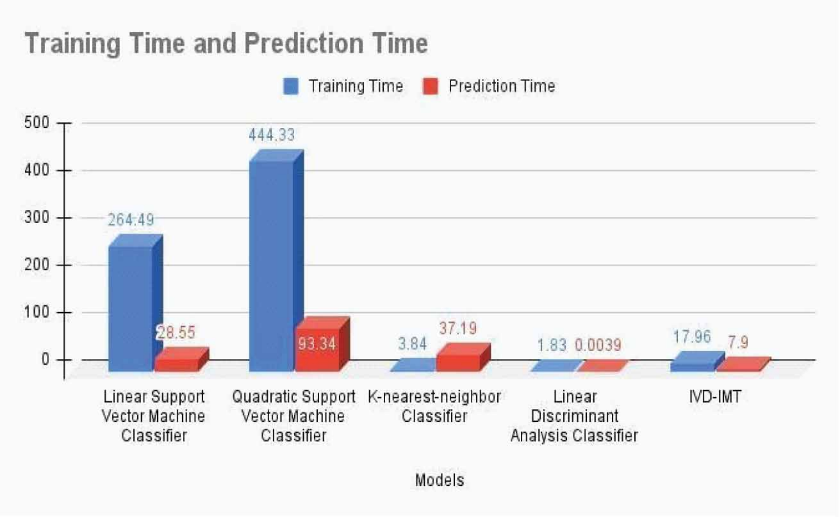


Figure 11b. Comparative Analysis of time complexity (in sec) with alternative models to detect intrusion



5.4 Discussion

After running several simulations on the varied datasets, anomaly detection via the IVD-IMT algorithm has delivered comparable detection rate and false alarm rates to neural networks for multi-dimensional data points, with improved time complexity over the conventional variable detector algorithm. This study also analyses the contemporary models regarding flexibility to new normals, computational complexity, and training time required before the systems are functional. The statistical nature of this algorithm gives it an edge over machine learning models that cannot be dynamically tuned without additional resource investment. IoE applications are not restricted to software units (Del Gaudio & Hirmer, 2021; Lata & Kumar, 2021; Some et al., 2021) that can run computationally complex algorithms, and each stage from the edge to the cloud needs to be secured. While the results for 2-dimensional data are in line with the research in Computer Immune systems, the model shows steep deviations when sample space goes beyond 3-dimensions. The concept of vicinity spheres is able to exploit spatial proximity to identify self and non-self entities on the X-Y-Z planes, but it opens the gates to a much deeper investigation when one needs to determine the anomaly based on 4 or more key attributes. Feature extraction through Principal Components Analysis has been pertinent to the dimensional limitation of this algorithm, but it's important to measure how it modifies the original data. The input features are transformations on real data and make it difficult to trace the outlying attribute for detected anomalies that may prove to be essential during mitigation.

5.5 Conclusion and Future Work

The proposed IoE-based algorithm dynamically allows two-way communication for detectors of variable size. These variable size detectors were tested on real-valued data points in multiple dimensions, for different values of self-threshold and their performance was analyzed. The detection rate was improved by 24.8% and the false alarm rate was found to be decreased by 27.69% when compared with the conventional Variable Detector algorithm. Comparative analysis with other machine learning models proved that the algorithm has comparable accuracy, or detection rate along with the lowest false alarm rate. The algorithm achieved coverage of 99%, with a threshold value of 0.01, i.e., closely covering the boundary of the self-spaces with this limit set at 1000 detectors. For estimating the area covered in the full non-self-region the algorithm V-detector utilizes the point percentage enclosed by the detectors. Overall, a detection rate of about 7.76% and a false alarm rate of 1.88% were achieved for 99% target coverage in the two-dimensional space, 97.22% in the three-dimensional space. From the experimental results, the proposed algorithm shows extremely promising results for handling two- dimensional and three-dimensional spaces and can be modelled for similar anomaly detection models with real-valued, multidimensional data that can be created from multiple features in a data-set. We intend to improve the algorithm accuracy for higher dimension values in the future. As with increasing dimensions the detection rate tends to fall whereas the false alarm rate is inclined towards the increase thereby reducing the accuracy.

REFERENCES

Angelov, P. P., Silva, G. C., & Dasgupta, D. (2016). *A survey of recent works in artificial immune systems. In Handbook on computational intelligence* (pp. 547–586). 005. World Scientific Publishing.

Balthrop, J., Esponda, F., Forrest, S., & Glickman, M. (2002). Coverage and generalisation in anartificial immune system. In Genetic and Evolutionary Computation, 3–10.

Banu, R., Ahammed, G. F. A., & Fathima, N. (2016). A review on biologically inspired approaches to security for Internet of things (IoT). In *International Conference on Electrical, Electronics, and Optimization Techniques(ICEEOT),* (pp. 1062–1066). IEEE. doi:10.1109/ICEEOT.2016.7754848

Barani, F. (2014). A hybrid approach for dynamic intrusion detection in adhoc networks usinggenetic algorithm and artificial immune system. In *Journal of Intelligent Systems (ICIS) Iranian Conference*, (pp. 1–6). IEEE.

Bayar, N., Darmoul, S., Hajri-Gabouj, S., & Pierreval, H. (2015, November). Fault detection, diagnosis and recovery using artificial immune systems: A review. *Engineering Applications of Artificial Intelligence*, *46*, 43–57. doi:10.1016/j.engappai.2015.08.006

Bendiab, E., & Kholladi, M.-K. (2010). The negative selection algorithm: A supervised learning approach for skin detection and classification. *IJCSNS International Journal of Computer Science and Network Security*, *10*, 86–92.

D'haeseleer, P., Forrest, S., & Helman, P. (1996). An immunological approach to change detection: Algorithms, analysis, and implications. In *Proceedings of the 1996 IEEE Symposium on Computer Security and Privacy*, (pp. 110–119). IEEE Computer Society. doi:10.1109/SECPRI.1996.502674

Dasgupta, D., Ji, Z., & Gonzalez, F. (2004). Artificial immune system (AIS) Research in the Last Five Years. *Proceedings*. (pp. 123–130). CEC.

Del Gaudio, D., & Hirmer, P. (2021). The IoT Vision: Challenges and Research Gaps. *International Journal of Organizational and Collective Intelligence*, *11*(4), 1–12. doi:10.4018/IJOCI.2021100101

Dong, B., Yang, J., Ma, Y., & Zhang, X. (2016). Medical monitoring model of Internet of things based on the adaptive threshold difference algorithm. *International Journal of Multimedia and Ubiquitous Engineering*, *11*(5), 75–82. doi:10.14257/ijmue.2016.11.5.08

Dwivedi, A., Srivastava, G., Dhar, S., & Singh, R. (2019). A Decentralized Privacy-Preserving Healthcare Blockchain for IoT. *Sensors (Basel)*, *19*(2), 326. doi:10.3390/s19020326 PMID:30650612

Foley, P. G. (2021). *Security Measures in IoT Devices, Including Wireless Medical Devices: Factors Influencing the Adoption of Effective Security Measures.* IGI Global. https://www.igi-global.com/article/security-measures-in-iot-devices-including-wireless-medical-devices/308267

Forrest, S., Perelson, A., Allen, L., & Cherukuri, R. (1994). Self-nonself discrimination in a computer. In *Proceedings of the 1994 IEEE Symposium on Research in Security and Privacy* (pp. 202–212). IEEE Computer Society Press. doi:10.1109/RISP.1994.296580

Hart, E. (2005). Not all balls are round: An investigation of alternative recognition-region shapes. In Lecture Notes in Computer Science, 29–42.

Hidayat, E., Fajrian, N. A., Muda, A. K., Huoy, C. Y., & Ahmad, S. (2011). A comparative study of feature extraction using PCA and LDA for face recognition. *7th International Conference on Information Assurance and Security (IAS),* (pp. 354–359). IEEE.

Igawa, K., & Ohashi, H. (2009). A negative selection algorithm for classification and reduction ofthe noise effect. Journal of Applied Soft Computing, 9, 431–438.

Ji, Z., & Dasgupta, D. (2004). Real-valued negative selection algorithm with Variable-sized detectors. *Lecture Notes in Computer Science, 30*, 287–298.

Ji, Z., & Dasgupta, D. (2005). Estimating the detector coverage in a negative selection algorithm. In H.-G. Beyer et al. (Eds.). GECCO 2005. *Proceedings of the 2005 Conference on Genetic and Evolutionary Computation*, 1 (pp. 281–288). ACM Press. doi:10.1145/1068009.1068056

Ji, Z., & Dasgupta, D. (2009). V-detector: An efficient negative selection algorithm with “probably adequate” detector coverage. *Information Sciences*, *179*(10), 1390–1406. doi:10.1016/j.ins.2008.12.015

Kumar, K. B. S., & Bairavi, K. (2016). IoT based health monitoring system for autistic patients. Proceedings of the Symposium on Big Data and Cloud Computing Challenges. *Smart Innovation, Systems and Technologies*. doi:10.1007/978-3-319-30348-2_32

Langley, D. J., van Doorn, J., Ng, I. C. L., Stieglitz, S., Lazovik, A., & Boonstra, A. (2021). The Internet of Everything: Smart things and their impact on business models. *Journal of Business Research*, *122*, 853–863. doi:10.1016/j.jbusres.2019.12.035

Lata, M., & Kumar, V. (2021). Standards and Regulatory Compliances for IoT Security. *International Journal of Service Science, Management, Engineering, and Technology*, *12*(5), 133–147. doi:10.4018/IJSSMET.2021090109

Lhazmir, S., El Moudden, I., & Kobbane, A. (2017). Feature extraction based on principal component analysis for text categorization. *International Conference on Performance Evaluation and Modeling in Wired and Wireless Networks (PEMWN),* (pp. 1–6). IEEE.

Meena, G., & Choudhary, R. R. (2017). A review paper on IDS classification using KDD 99 and NSL KDD dataset in Weka. *International Conference on Computer, Communications and Electronics (Comptelix),* (pp. 553–558). . 8004032 IEEE.

Miraz, M., Ali, M., Excell, P., & Picking, R. (2018). Internet of nano-things, things and everything: Future growth trends. *Future Internet, 10*(8), 68.

Muhammad, G., Rahman, S. M. M., Alelaiwi, A., & Alamri, A. (2017, January). AL elaiwi, and A. Alamri. *IEEE Communications Magazine*, *55*(1), 69–73. doi:10.1109/MCOM.2017.1600425CM

Murali, M. (2015). Principal component analysis based Feature Vector Extraction. *Indian Journal of Science and Technology*, *8*(35). Advance online publication. doi:10.17485/ijst/2015/v8i35/77760

Pamukov, M. E., & Poulkov, V. K. (2017). Multiple negative selection algorithm: Improving detection error rates in IoT intrusion detection systems. *9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS),* (pp. 543–547). IEEE.

Patel, K. K., & Patel, S. M. (2016). Internet of things IOT: Definition, characteristics, architecture, enabling technologies, application future challenges. *International Journal of Engineering Science and Computing*, *6*(5), 6122–6131.

Read, M., Andrews, P., & Timmis, J. (2012). *An introduction to Artificial Immune systems*. doi: 47.10.1007/978-3-540-92910-9

Ryan, P. J., & Watson, R. B. (2017). Research challenges for the Internet of things: What role can OR play? *Systems*, *5*(1), 1–34. doi:10.3390/systems5010024

Shapiro, J. M., Lamont, G. B., & Peterson, G. L. (2005). An evolutionary algorithm to generate hyperellipsoid detectors for negative selection. In H.-G. Beyer et al. (Eds.) *Proceedings of the 2005 Conference on Genetic and Evolutionary Computation,* 1 (pp. 337–344). ACM Press.

Some, E., Boots, B., & Gondwe, G. (2021). Enabling 5G and IoT: In Search of More Spectrum for Connected Devices. *International Journal of Interdisciplinary Telecommunications and Networking*, *13*(4), 1–10. doi:10.4018/IJITN.2021100101

Srivastava, G., & Lin, J. (2021). Chun-Wei and Pirouz, matin and Li, Yuanfa and Yun,Unil, A Pre-Large Weighted-Fusion System of Sensed High-Utility Pattern. *IEEE Sensors Journal*. doi:10.1109/JSEN.2020.2991045

Sunke, B. (2008). *Research and analysis of network intrusion detection system* [Thesis]. GTBIT, GGSIPU, India.

Tan, Y. (2016). *Artificial immune system: Applications in. Computers and Security, 1*(18). Wiley. https://www. wiley. com/en-us/Artificial

Thirumalai, C., & Mohan, S., & Srivastava, G. (2020). An efficient public, key secure scheme for cloud and IoT security. In Computer Communications.

Tiwari, M., Kumar, R., Bharti, A., & Kishan, J. (2017). Intrusion Detection System. *International Journal of Technical Research and Applications.*, *5*, 2320–8163.

Xu, J., Wang, J., Xie, S., Chen, W., & Kim, J. (2013) January,. Study on intrusion detection policy for Wireless sensor networks. *International Journal of Security and its Applications, 7*(1), 1–6.

Yang, C., Jia, L., Chen, B., & Wen, H. (2020). Negative selection algorithm based on AntigenDensity clustering. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 44967–44975. doi:10.1109/ACCESS.2020.2976875

Yang, H., Li, T., Hu, X., Wang, F., & Zou, Y. (2014). A survey of artificial immune systembased intrusion detection. *TheScientificWorldJournal*, 2014.

Yeole, A. S., & Kalbande, D. R. (2016). Use of Internet of Things (IoT) in Healthcare: A Survey. In *Proceedings of the ACM Symposium on Women in Research 2016 (WIR '16)* (pp. 71–76). Association for Computing Machinery. doi:10.1145/2909067.2909079

Parul Lakhotia received her Bachelor of Engineering degree in Information Technology from the Netaji Subhas University of Technology in 2021. She's working as an IT professional in the R&D department at Salesforce. She's been researching in the field of Network Security, IoT, and fog computing.

Rinky Dwivedi has completed her B.Tech in Computer Science and Engineering from Guru Gobind Singh Indraprastha University, Delhi in 2004 and M.E. in Computer Technology and Application from Delhi College of Engineering, Delhi in 2008. She received her Doctorate in 2016 from Delhi Technological University, New Delhi . Dr. Rinky has over 19 years of experience in Academics, currently working as Associate Professor and Head of the Department in Computer Science Engineering Department of Maharaja Surajmal Institute of Technology, New Delhi, INDIA. She has published more than 20 research papers in reputed Journals and conferences proceedings and has also authored books.

Deepak Kumar Sharma is working as an Associate Professor in the Department of Information Technology, Indira Gandhi Delhi Technical University for Women (IGDTUW), Kashmere Gate, Delhi, India. Earlier, he has worked as Assistant Professor at Netaji Subhas University of Technology (Formerly N.S.I.T.), Dwarka, Delhi. He obtained his Ph. D in Computer Engineering from University of Delhi, India in 2016. His research interests include opportunistic networks, wireless ad hoc and sensor networks, Software Defined Networks and IoT Networks. He has over 17 years of experience in Academics. He has published various research papers in reputed international journals like ETT Wiley, IEEE Systems Journal, IEEE IoT Journal, Computer Communication Elsevier, IJCS Wiley etc. and conferences of repute like IEEE AINA, GLOBECOM etc. He has also authored various book chapters in edited books of IET, Wiley, Springer, Elsevier etc. He has served as session chair in many conferences and is also a reviewer of various reputed journals like ETT Wiley, AIHC Springer, IJCS Wiley etc.

Nonita Sharma is working as Associate Professor, Indira Gandhi Delhi Technical University for Women, New Delhi. She has more than 15 years of teaching experience. Her major area of interest includes data mining, bioinformatics, time series forecasting and wireless sensor networks She has published several papers in the International/National Journals/Conferences and book chapters. She received best paper award for her research paper in Mid-Term Symposium organized by CSIR, Chandigarh. She has been awarded Best Teacher Award in view of recognition of contributions, achievements, and excellence in Computer Science & Engineering in NIT Jalandhar. She has been awarded Best Content Guru Award by Infosys twice. She has authored a book titled- "Analysis of Algorithms". She has been the editor of various books published by eminent publishers like WILEY, Taylor & Francis, CRC Press etc. She is member, IEEE and has been shortlisted in Top 5 for IEEE Women Achiever Award. She is the reviewer of many peer reviewed journals and contributed to academic research in terms of projects, papers, and patents.