Computer Systems Science & Engineering
DOI: 10.32604/csse.2023.029680
Article

Tech Science Press

A Novel Approach for Network Vulnerability Analysis in IIoT

K. Sudhakar* **and S. Senthilkumar**

Department of Computer Science and Engineering, University College of Engineering, Pattukkottai, Rajamadam, 614701, India
*Corresponding Author: K. Sudhakar. Email: ksudhakar.cs@gmail.com
Received: 09 March 2022; Accepted: 11 April 2022

Abstract: Industrial Internet of Things (IIoT) offers efficient communication among business partners and customers. With an enlargement of IoT tools connected through the internet, the ability of web traffic gets increased. Due to the raise in the size of network traffic, discovery of attacks in IIoT and malicious traffic in the early stages is a very demanding issues. A novel technique called Maximum Posterior Dichotomous Quadratic Discriminant Jaccardized Rocchio Emphasis Boost Classification (MPDQDJREBC) is introduced for accurate attack detection with minimum time consumption in IIoT. The proposed MPDQDJREBC technique includes feature selection and categorization. First, the network traffic features are collected from the dataset. Then applying the Maximum Posterior Dichotomous Quadratic Discriminant analysis to find the significant features for accurate classification and minimize the time consumption. After the significant features selection, classification is performed using the Jaccardized Rocchio Emphasis Boost technique. Jaccardized Rocchio Emphasis Boost Classification technique combines the weak learner result into strong output. Jaccardized Rocchio classification technique is considered as the weak learners to identify the normal and attack. Thus, proposed MPDQDJREBC technique gives strong classification results through lessening the quadratic error. This assists for proposed MPDQDJREBC technique to get better the accuracy for attack detection with reduced time usage. Experimental assessment is carried out with UNSW_NB15 Dataset using different factors such as accuracy, precision, recall, F-measure and attack detection time. The observed results exhibit the MPDQDJREBC technique provides higher accuracy and lesser time consumption than the conventional techniques.

Keywords: Industrial internet of things (iiot); attack detection; features selection; maximum posterior dichotomous quadratic discriminant analysis; jaccardized rocchio emphasis boost classification

1 Introduction

Industrial IoT (IIoT) signifies the function of IoT in industrial organization to get better operational efficiency. The IIoT speeds up the industrial mechanization development through connecting many IoT devices. The extensive exploitation of IoT apparatus in the Industrial model has generated diverse

applications i.e., smart manufacturing, intelligent transport, etc. However, IIoT applications are impacted by dangerous security hazards caused by means of anomalous IIoT devices. Machine learning techniques are more efficient for securing the IIoT applications from harsh attacks.

A new Hybrid Deep Random Neural Network (HDRaNN) was developed in [1] for cyberattack detection. The designed approach increases the accuracy of attack detection however the time complexity was more. A hinge classification algorithm was introduced in [2] to detect the various attacks. The designed technique increases the precision but it failed to apply the large volume of data samples in a wide range of scenarios. Different machine learning techniques were introduced in [3] for cyber-vulnerability assessment. In [4], a machine learning technique was developed for distinguishing the data into different classes. However, it failed to reduce the complexity.

A False Data Injection" (FDI) attack detection using Autoencoders was developed in [5] for discovering attacks with lesser time. But it failed to use the efficient dimensionality reduction method to further minimize the execution time.

A Distributed Congestion Control system was introduced in [6] to identify and mitigate DoS attacks. But it failed to examine the traffic segmentation for improving detection and mitigation attack. A novel neural network approach was developed in [7] for the recognition of various anomalies in Industrial IoT systems. However, it failed to further validate the accuracy of attack detection.

Machine learning algorithms were developed [8] for anomaly detection in an industrial control environment. But the accuracy investigation was not performed. Deep Reinforcement Learning was planned in [9] for finding attacks in IIoT. The designed learning system was not efficient to achieve higher attack detection accuracy.

A novel IIoT-based dataset, called TON_IoT was developed in [10] that integrate both normal sensor measurement data and a variety of attacks. However, false positive rate was higher.

The key contributions of the MPDQDJREBC technique are listed as follows,

- A novel MPDQDJREBC technique is developed for identifying malicious attacks with the applicability of the machine learning ensemble technique.
- To get better malicious attacks detection performance in IIoT communication network, the MPDQDJREBC technique is developed based on feature extraction and categorization.
- To minimize the attack detection time, a Maximum Posterior Dichotomous Quadratic Discriminant analysis is performed in the MPDQDJREBC technique to find the significant features for accurate classification based on likelihood estimation.
- After that, ensemble classification is performed using the Jaccardized Rocchio Emphasis Boost technique to combine the weak learner result into strong output and also minimize the quadratic error. This assists to get better true positives and diminish the false positives and false negatives.
- Finally, well-known experimentation is carried out to measure the performance of our MPDQDJREBC technique and other existing works. The experimental result reveals that MPDQDJREBC technique is highly efficient for network vulnerability analysis in IIoT.

This paper is intended as below. Section 2 explains conventional studies using machine learning techniques. Section 3 shows detailed processes of proposed MPDQDJREBC technique. In Section 4, experimental evaluation and results of these methods are presented. The conclusion Section 5 depicts the results of proposed work.

2 Related Works

A novel multi-feature layer was designed in [11] for attack detection where it gets lesser the false positive rate. But the time utilized for attack detection was more. A principal component analysis (PCA) was introduced in [12] for predicting malicious activities in IIoT. However, accurate detection of abnormal events was not attained. A Deep-IFS was developed [13] to discover attacks in IIoT traffic. But the efficient dimensionality reduction technique was not applied to minimize the complexity of intrusion detection.

A process-level attack-detection method was introduced in [14] for capable of detecting stealthy attacks. But it failed to apply the efficient machine learning technique for improving the accuracy of attack detection. A deep learning-based IDS method was developed in [15] for IIoT. However, securing IIoT network from various kinds of malicious activities was considered.

Malware attacks detection was performed in [16] for IIoT. However, accuracy measured during attack discovery was not higher. A novel Convolutional Neural Network-Long Short Term Memory (AMCNN-LSTM) system was introduced in [17] for distinguishing the anomalies. However, the designed framework was not robust for accurate anomaly detection models. A novel classification scheme was introduced in [18] for the traffic characterization mechanism. However, the different levels of features were not considered to improve the performance of classification.

A novel malware discovery method using a machine-learning approach (MADP-IIME) was introduced in [19]. However, it failed to consider diverse categories of malware attacks. A graph-based security framework was developed in [20] for IIoT network. However, the attack detection performance using this framework was poor.

3 Proposal Methodology

The IIoT includes number of sensors, apparatus, engineering applications, databases, services, etc. The sensors employed in the IIoT network generate a massive amount of information where discovering the behavior of the network is considered as vital for securing IIoT applications from attacks. Therefore, an efficient attack recognition system is necessitated to safeguard an IIoT. Based on the motivation, the MPDQDJREBC technique is introduced for cyber-vulnerability assessment through the attack detection with different processes namely feature selection and classification.

Fig. 1 demonstrates the overall process of MPDQDJREBC for accurate attack detection. The proposed MPDQDJREBC technique consists of two phases. First, the number of features and data are collected from the dataset. After that, the feature selection is performed using Maximum Posterior Dichotomous Quadratic Discriminant analysis. Then, a classification is done based on the feature Jaccardized Rocchio Emphasis Boosting technique. A detailed explanation of these two processes of the MPDQDJREBC technique is provided in the below sections.

3.1 Maximum Posterior Dichotomous Quadratic Discriminant Analysis

The Maximum Posterior Dichotomous Quadratic Discriminant Analysis (MPDQDA) is introduced in this research work to detect the significant features. MPDQDA is a machine learning dimensionality reduction technique that helps to analyze the features in the given dataset. While using a machine learning technique for IoT real-time dataset, a lot of features are presented in the dataset and not all these features are significant every time. This unnecessary feature while training the machine learning technique directs to decrease the overall accuracy and enhance the complexity. Therefore, a feature selection is one of the significant steps while making a machine learning technique to discover the best possible set of features.

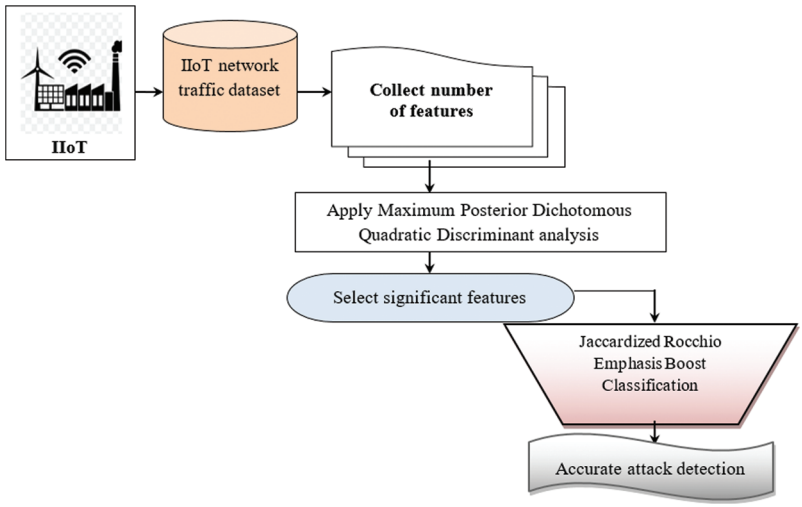

Figure 1: The processing diagram of proposed MPDQDJREBC technique

Fig. 2 illustrates the block diagram of MPDQDA to find the significant features in the dataset. First, number of features '$a_1, a_2, a_3, \ldots a_n$' are collected from the dataset. After that, Quadratic discriminant analysis is performed based on the Gaussian distribution.

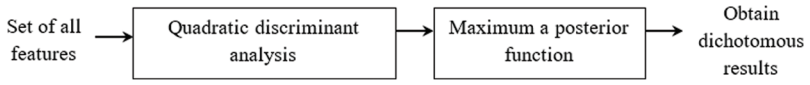

Figure 2: The block diagram of the MPDQDA

$$\delta\left(a_i, o_j\right) = \frac{1}{\sqrt{2\pi}\varphi}\exp\left[-0.5\left(\frac{a_i - o_j}{\varphi}\right)^2\right] \tag{1}$$

where, $\delta\left(a_i, o_j\right)$ represents a likelihood between the features, φ represents the deviation, a_i, o_j represents features and objectives (*i.e.*, attack detection) in the dataset. Then the Maximum posterior function is applied to find the maximum likelihood between the features.

$$\rho_{map} = arg\ max\delta\left(a_i, o_j\right) \tag{2}$$

where, ρ_{map} Maximum a posterior function, arg max denotes an argument of the maximum function, $\delta\left(a_i, o_j\right)$ indicates likelihood between the features and objective. Finally, the dichotomous outputs are evaluated according to the likelihood estimation. Dichotomous outputs precisely provide two distinct outcomes.

$$y = \begin{cases} arg\ max\delta\left(a_i, o_j\right); & Select\ features \\ Otherwise; & Remove\ features \end{cases} \tag{3}$$

From (3), 'y' denotes Dichotomous results. The feature which is more likelihood is selected as significant for further processing. Otherwise, the features are removed from the dataset. As a result, the time complexity of attack detection in the IIoT network is said to be minimized.

The algorithmic process of MPDQDA is described as given below,

Algorithm 1: Maximum posterior dichotomous quadratic discriminant analysis based feature selection

Input: Dataset D, features or attributes $a_i = \{a_1, a_2, a_3, \ldots a_n\}$, data $D_1, D_2, D_3, \ldots D_n$
Output: select significant features
Begin
Collect the number of features or attributes $a_1, a_2, a_3, \ldots a_n$
For each feature 'a_i'
Measure likelihood function '$\delta\ (a_i, o_j)$'
Apply Maximum a posterior 'ρ_{map}'
if $(\arg\max \delta\ (a_i, o_j))$ then
Select significant feature
else
Remove the feature
end if
Return (significant features)
end for
end

Algorithm 1 explains the algorithmic process of Maximum posterior dichotomous quadratic discriminant analysis based feature selection. At first, collections of features are taken from the dataset as input. Then the likelihood is estimated between the features and objective functions. Followed by, maximum a posterior is used to find the maximum likelihood between the features. Based on maximum likelihood estimation, vital features are discovered and other irrelevant features are eliminated which resulting in reduced time consumption of the attack detection.

3.2 Jaccardized Rocchio Emphasis Boosting Based Attack Detection

After the feature selection, the proposed MPDQDJREBC technique predicts malicious attacks with the application of machine learning concepts. The malicious attacks in IIoT systems are discovered in this section using Jaccardized Rocchio Emphasis boosting based attack detection (JREB-AD). In proposed work, JREB-AD is a machine learning ensemble technique. JREB-AD translates weak learners into strong ones. The MPDQDJREBC technique uses the Jaccardized Rocchio Emphasis Boost method for accurate vulnerability assessment in IIoT systems by identifying malicious attacks.

Fig. 3 explains schematic structure of JREB-AD for accurate classification with lesser time consumption. The JREB-AD acquires the training sample set $\{x_i, z_i\}$ as input where $x_i = D_1,\ D_2,\ \ldots, D_m$' denotes the sample data and z_i indicates the ensemble classification outcomes. Then, JREB-AD builds 'k' set of weak learners $\beta_1, \beta_2, \beta_3, \ldots . \beta_k$ as a Jaccardized Rocchio Classifier (JRC) to categorize input data $D_1, D_2, D_3, \ldots D_m$ into different classes.

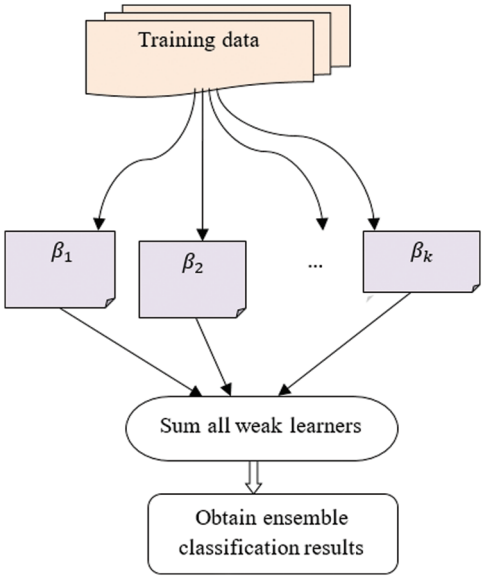

Figure 3: Schematic structure of JREB-AD

The JRC is a classification model that assigns the input training samples into the label of the output class whose mean is closest to the observation. Here, two classes are initialized namely normal and attack. The Jaccard index is applied to a Rocchio classifier to measure the similarity between the training data and mean of class.

$$S = \frac{[D_i \cap C_m]}{\sum D_i + \sum C_m - [D_i \cap C_m]} \tag{4}$$

where 'S' indicates a Jaccard similarity coefficient, D_i denotes training data, C_m indicates a mean of a particular class, the intersection symbol '$\cap$' designates mutual independence between the data and mean of class which are statistically dependent, $\sum D_i$ is the summation of D_i value, $\sum C_m$ is the summation of C_m value. The similarity coefficient provides results from 0 to 1 $[0 \leq S \leq 1]$. The proposed JREB-AD classifies the data into a particular class based on the higher similarity. In this way, the weak JRC classifies the data into normal or attack.

The observed weak JRC results have some training errors. The JREB-AD sum the entire weak JRC results using below,

$$z = \sum_{i=1}^{k} R_i \tag{5}$$

where, z indicates ensemble classification results, R_i indicates weak JRC result. The weight is initialized for all weak JRC to determine the final strong classification results using below,

$$z = \sum_{i=1}^{k} R_i * \varphi_i \tag{6}$$

From (6), 'φ_i' depicts weight given to the weak JRC. The JREB-AD employs the weighted emphasis function to measure the quadratic error of classification results using below,

$$\epsilon = \exp\left[\vartheta\left(\left(\sum_{i=1}^{k} R_i\varphi_i \; - z\right)^2 - (1-\vartheta)\left(\sum_{i=1}^{b} R_i\right)^2\right)\right] \tag{7}$$

From (7), ϵ shows a weighted emphasis function, ϑ indicates a weighting constraint, z depicts actual results, '$\sum_{i=1}^{k} R_i\varphi_i$' indicates a predicted classification result of weak JRC with weight φ_i and without weight $\sum_{i=1}^{k} R_i$. From (7), ϑ denotes a weighting constraint rate. The JREB-AD obtain the final output using below,

$$\epsilon = \exp\left[\left(\sum_{i=1}^{k} R_i\varphi_i - z\right)^2\right] \tag{8}$$

According to estimate the error value, the weak JRC weight gets updated. If the weak JRC correctly classifies input information, the weight value is decreased. Otherwise, the weight is increased. Finally, the weak JRC with lowest error is considered as the final strong result for accurate classification. Based on the classification results, normal or attack data are correctly identified with superior precision.

The detailed process of JREB-AD is illustrated as follows,

Algorithm 2: Jaccardized Rocchio Emphasis Boosting Based Attack Detection

Input: Selected features, data $D_1, D_2, D_3, \ldots D_n$

Output: Improve the accuracy

Begin

1. For each data 'D_i'
2. Construct 'k' number of weak learners
3. Initialize the classes and mean
4. Measure the Jaccard similarity 'S'
5. Classifies the data into a particular class
6. end for
7. Combine all weak learner results '$z = \sum_{i=1}^{k} R_i$'
8. for each weak classifier results
9. Assign the weight ' φ_i'
10. evaluate the quadratic error ' ϵ'
11. Update the weight ' $\nabla\varphi_i$'
12. Find the weak JRC with lowest error
18: Return (accurate classification output)
19: end for

End

Algorithm 2 represents overall process of JREB-AD to get better the categorization precision. The JREB-AD at the start builds a set of weak JRC. The weak JRC initializes the number of classes and means value. Then the Jaccard similarity is measured between the data and mean of a particular class. Based on similarity, the information is categorized into a consequent class. The obtained weak learner results are combined and assigned weight. The weighted emphasis function is employed in JREB-AD to calculate the quadratic error of each weak JRC. Finally, the weak JRC with a lesser quadratic error is identified as an accurate strong classification result.

4 Experimental Scenario

Performance evaluations of the proposed MPDQDJREBC technique and existing methods namely HDRaNN [1], HCA-MBGDALRM [2] are measured by conducting experimental using Java language and UNSW_NB15 Dataset from Kaggle https://www.kaggle.com/mrwellsdavid/unsw-nb15. In order to conduct the experiment, a UNSW_NB15 Dataset is used for detecting the normal or attacks. The dataset consists of different. CSV files. Among them, the training.CSV files are taken to conduct the experiments. The training.CSV files consist of 1, 75,341 records and 45 attributes. The last two columns of the datasets are attack category and Label. Each row in the dataset is classified and labeled as normal or attack records. Before the data classification, significant features are selected for minimizing the complexity. The result of MPDQDJREBC technique is estimated using below metrics.

- Accuracy
- Precision
- Recall
- F-measure
- Attack detection time

Accuracy: It is calculated as ratios of a number of exactly predicted data to the total number of data considered as input Then, the accuracy is obtained as,

$$Acc = \left(\frac{pos_t + neg_t}{pos_t + neg_t + pos_f + neg_f} \right) * 100 \tag{9}$$

where Acc denotes an accuracy that is evaluated in units of percentage (%).Where, pos_t denotes a true positive, neg_t indicates a true negative, pos_f denotes a false positive, neg_f denotes a false negative. Accuracy is measured in milliseconds (ms).

Precision: It is obtained as the ratio of the number of precisely classified data to the total number of input data.

$$Pr = \left[\frac{pos_t}{pos_t + pos_f} \right] * 100 \tag{10}$$

where, Pr denotes a precision, pos_t denotes a true positive, pos_f indicates a false positive. Precision is measured in percentage (%).

Recall: It is acquired as a percentage of true positives to true positive and false negatives. The recall is calculated as follows,

$$Rl = \left[\frac{pos_t}{pos_t + neg_f} \right] * 100 \tag{11}$$

where, Rl denotes a recall, pos_t denotes a true positive, neg_f denotes a false negative. The recall is measured in percentage (%).

F-measure: it is measured based on the average of precision and recall. The F-measure is computed as given below,

$$Measure_f = 2*\left[\frac{Pr*Rl}{Pr+Rl}\right]*100 \tag{12}$$

The F-measure is calculated in units of percentage (%).

Attack detection time: It calculates time required by the algorithm for detecting the attack or normal through the classification process. The attack detection time is calculated as given below,

$$T = m*T\ (CSD) \tag{13}$$

Where T indicates an attack detection time, m denotes the number of data, $T\ (CSD)$ explains a time employed to categorize the single data. The attack detection time is obtained in units of milliseconds (ms).

Tab. 1 displays the accuracy results of three different machine learning methods namely MPDQDJREBC, HDRaNN [1], and HCA-MBGDALRM [2] respectively. As shown in Tab. 1, ten dissimilar accuracy results are observed based on number of input data. The result clearly illustrates that the proposed MPDQDJREBC achieves higher accuracy when compared to existing techniques. In the first iteration, 5000 data are considered to calculate the accuracy. By applying MPDQDJREBC, 92% accuracy was observed and the accuracy of existing HDRaNN [1] and HCA-MBGDALRM [2] are 88% and 84% respectively. Similarly, nine various performance results are observed for each technique. The results point out that the accuracy using proposed MPDQDJREBC technique is improved by 4%, 6% as compared to HDRaNN [1] and HCA-MBGDALRM [2] respectively.

Table 1: Comparison of accuracy

Number of data	Accuracy (%)		
	MPDQDJREBC	HDRaNN	HCA-MBGDALRM
5000	92	88	84
10000	94	90	88
15000	94.33	88.66	86.66
20000	94.5	90.5	89
25000	94.4	91.2	89.6
30000	94.33	91.66	90.47
35000	93.71	91.42	89.42
40000	94	92.25	90.5
45000	94.22	91.55	90.22
50000	94.4	91.6	90

Fig. 4 illustrates the accuracy of graphical representation with varying numbers of data. As shown in the figure, the horizontal axis signifies the number of data and the vertical axis denotes the accuracy of attack detection. Therefore, the graphical outcome clearly shows that the MPDQDJREBC technique attains higher accuracy than the conventional techniques. This is because of the fact that the application of Jaccardized Rocchio Emphasis Boost Classification technique. The ensemble technique accurately gives strong classification results through lessening the quadratic error. This supports to get better the accuracy of attack detection.

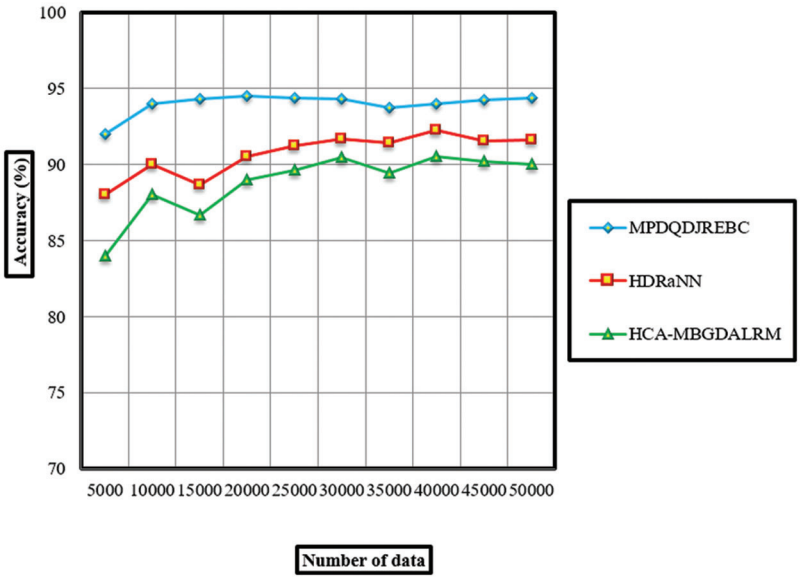

Figure 4: Graphical performance of accuracy

Tab. 2 and Fig. 5 reveal the performance results of precision with varying numbers of data. For each iteration, different precision results are observed for each method. The precision of three methods MPDQDJREBC, HDRaNN [1], and HCA-MBGDALRM [2] are shown in the above graph. The graphical outcomes inferred that the precision of the MPDQDJREBC technique is increased when the conventional works. This is owing to the employment of the proposed concepts accurately finding the normal or attack and minimizing the incorrect classification through the ensemble technique. For each technique, ten various results are estimated with respect to dissimilar numbers of input data. The obtained results of precision using the MPDQDJREBC technique are analyzed to the performance of traditional works. The compared results indicate that the performance of precision is found that the precision is considerably increased by 3% and 4% using MPDQDJREBC technique than the state-of-the-art methods.

Table 2: Comparison of precision

Number of data	Precision (%)		
	MPDQDJREBC	HDRaNN	HCA-MBGDALRM
5000	93.33	90.69	87.5
10000	95.69	93.25	91.95
15000	95.72	90.90	89.84
20000	96.27	92.77	92.04
25000	96.58	93.83	92.82
30000	96.44	94.22	93.28
35000	96.33	94.42	93.08
40000	96.53	95.14	94
45000	96.69	94.68	94.11
50000	96.80	94.80	93.84

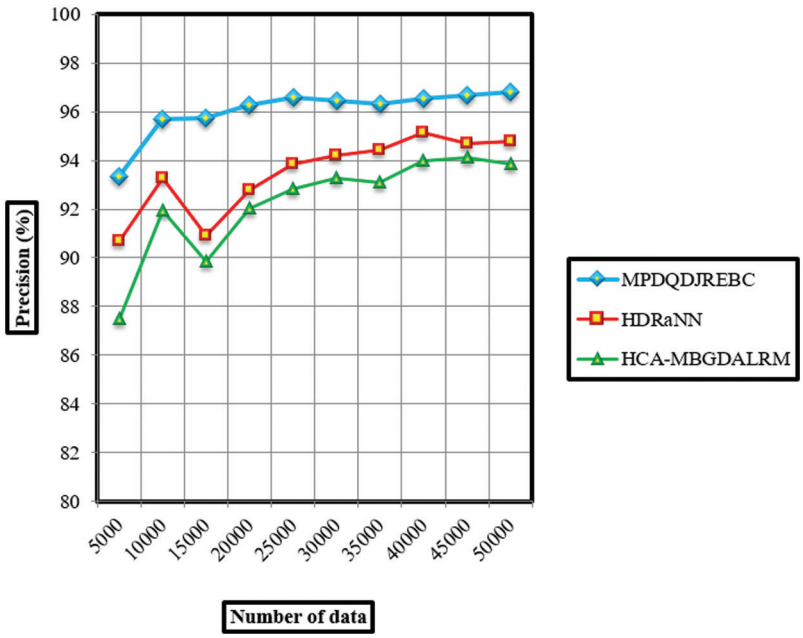

Figure 5: Performance of precision with varying number of data

The experimental results of recall using three methods are shown in Tab. 3 and Fig. 6. The recall is measured based on true positives and false-negative results of attack detection. The recall is measured based on three methods MPDQDJREBC, HDRaNN [1], and HCA-MBGDALRM [2]. Compared to existing methods, the MPDQDJREBC technique has the ability to provide a higher recall rate. This is due to analyzing the extracted feature values with the help of Jaccard similarity and it provides either two possible results such as normal or attack. When getting the 5000 data as input, the resultant value of the recall rate is 97.67%. Whereas, the recall value of the two existing methods HDRaNN [1] and HCA-MBGDALRM [2] are 95.12% and 92.10% respectively. In the same way, a variety of results are determined based on number of input data. The average of comparative analysis proves that the recall is found to be improved by 1% and 3% using the MPDQDJREBC model than the existing [1] and [2].

Table 3: Comparison of recall

Number of data	Recall (%)		
	MPDQDJREBC	HDRaNN	HCA-MBGDALRM
5000	97.67	95.12	92.10
10000	97.80	95.40	94.11
15000	98.17	96	94.26
20000	97.83	96.53	95.29
25000	97.41	96.38	95.39
30000	97.48	9	96.15
35000	96.92	96.21	95.17
40000	97.05	96.44	95.56
45000	97.14	96.07	95.04
50000	97.22	96.05	95.10

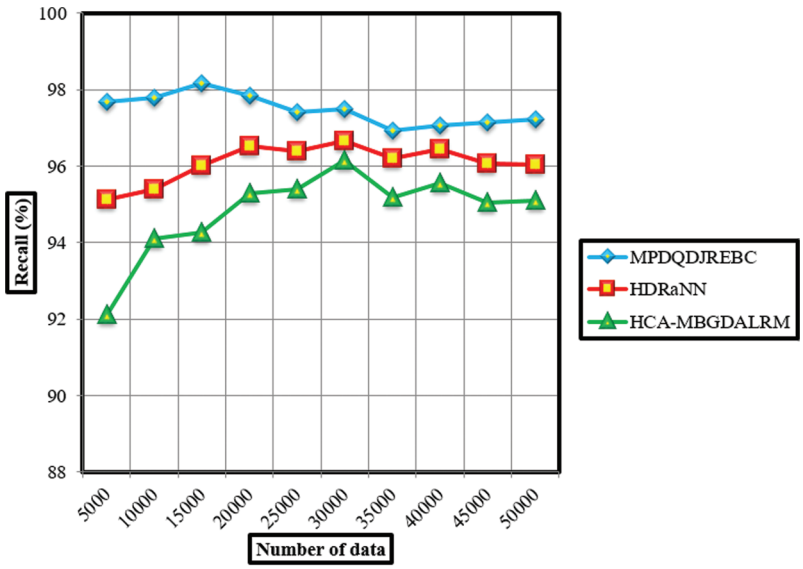

Figure 6: Performance of recall with varying number of data

Tab. 4 and Fig. 7 demonstrate the experimental results of the F-measure of attack detection using three methods MPDQDJREBC, HDRaNN [1], and HCA-MBGDALRM [2] depends on varied number of data. The F-measure is measured by both precision results and recall. From the tabulated value, it is demonstrated that the proposed model provides improved F-measure results when compared to existing techniques. Let us consider the 5000 input data, the F-measure of the MPDQDJREBC technique is 95.45% and the F-measure of HDRaNN [1] and HCA-MBGDALRM [2] are 92.85% and 89.74%. The compared results indicate that the MPDQDJREBC technique increases the performance of the F-measure by 2% and 3% when compared to [1] and [2] respectively.

Table 4: Comparison of F-measure

Number of data	F-measure (%)		
	MPDQDJREBC	HDRaNN	HCA-MBGDALRM
5000	95.45	92.85	89.74
10000	96.73	94.31	93.01
15000	96.92	93.38	91.99
20000	97.04	94.61	93.63
25000	96.99	95.08	94.08
30000	96.95	95.42	94.69
35000	96.62	95.30	94.11
40000	96.78	95.78	94.77
45000	96.91	95.36	94.57
50000	97	95.42	94.46

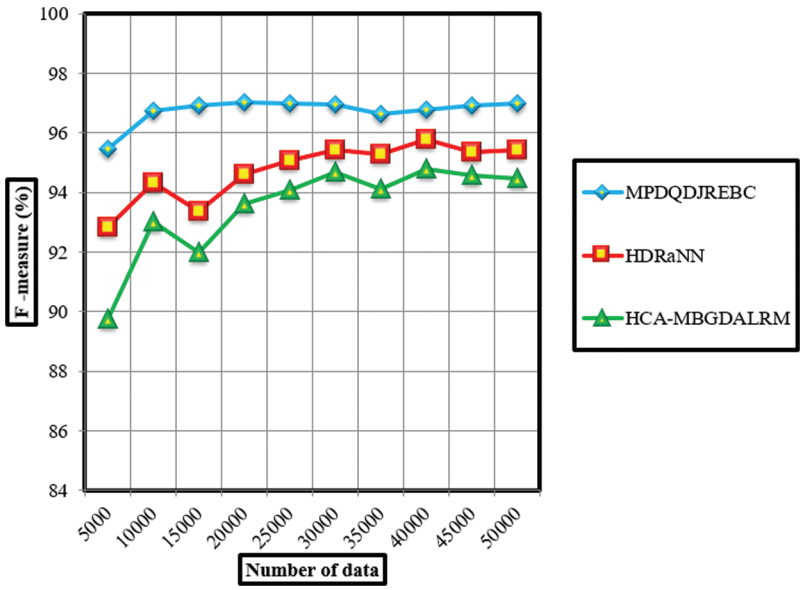

Figure 7: Performance of F-measure with varying number of data

Tab. 5 provides the performance results of attack detection time of MPDQDJREBC, HDRaNN [1] and HCA-MBGDALRM [2]. For each technique, ten various results are observed by considering a varied number of input network data taken from given dataset. The experimental output indicates that attack detection time of proposed MPDQDJREBC technique is minimized when compared to two conventional works. While acquiring 5000 data for estimating the attack detection time, the proposed MPDQDJREBC consumes 17.5 *ms* to classify an input data as normal or attack. Whereas, 19 *ms and* 21 *ms* of time consumed by the HDRaNN [1] and HCA-MBGDALRM [2] for classifying the input data. From the comparative analysis, it is clear that the attack discovery time using proposed MPDQDJREBC technique is lessened by 12% and 18% when compared to [1] and [2] respectively.

Table 5: Comparison of attack detection time

Number of data	Attack detection time (ms)		
	MPDQDJREBC	HDRaNN	HCA-MBGDALRM
5000	17.5	19	21
10000	20	25	27
15000	22.5	27	30
20000	26	30	32
25000	30	32.5	35
30000	33	36	39
35000	35	38.5	42
40000	36.8	42	44
45000	39.6	46.3	48.1
50000	41.5	47.5	49

In Fig. 8, the experimental results of attack detection time with the different numbers of data using three different techniques MPDQDJREBC, HDRaNN [1], and HCA-MBGDALRM [2]. Among the three methods, the attack detection time of MPDQDJREBC is considerably reduced. This is due to the application of MPDQDA based feature selection. The MPDQDA measures the likelihood estimation between the features and objective functions. The maximum posterior function helps to discover the significant features and other features are taken away from the dataset. This assists to diminish the time consumption of the attack detection.

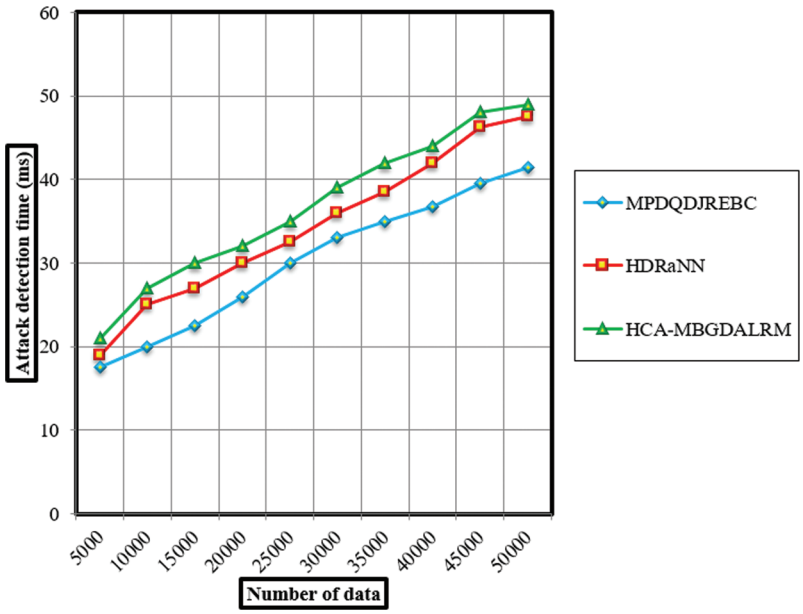

Figure 8: Performance of attack detection time with varying number of data

5 Conclusion

A novel machine learning-based technique MPDQDJREBC is intended in this article to efficiently discover of attacks in IIoT. The MPDQDJREBC technique cyber-vulnerability assessment along with their security susceptibilities through the accurately and timely detecting attacks. At first, the Maximum posterior dichotomous quadratic discriminant analysis is performed to emphasize the vital features. With the selected significant features, the Jaccardized Rocchio Emphasis Boost Classification technique is applied for detecting the normal or attack with higher accuracy. Experimental assessment is carried out using an improved version of the UNSW_NB15 Dataset with the parameters i.e., accuracy, precision, recall, F-measure, and attack discovery time. The quantitatively analyzed result clearly indicates that the MPDQDJREBC technique improves the accuracy, precision, recall, F-measure and also lessens attack detection time as than the conventional works.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] Z. E. Huma, S. Latif, J. Ahmad, Z. Idrees, A. Ibrar *et al.,* "A hybrid deep random neural network for cyberattack detection in the industrial internet of things," *IEEE Access*, vol. 9, pp. 55595–55605, 2021.

[2] X. Yan, Y. Xu, X. Xing, B. Cui, Z. Guo *et al.,* "Trustworthy network anomaly detection based on an adaptive learning rate and momentum in IIoT," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6182–6192, 2020.

[3] M. Zolanvari, M. A. Teixeira, L. Gupta, K. M. Khan and R. Jain, "Machine learning-based network vulnerability analysis of industrial internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6822–6834, 2019.

[4] D. J. Atul, R. Kamalraj, G. Ramesh, K. S. Sankaran, S. Sharma *et al.,* "A machine learning based IoT for providing an intrusion detection system for security," *Microprocessors and Microsystems*, vol. 82, no. 4, pp. 103741, 2021.

[5] M. M. N. Aboelwafa, K. G. Seddik, M. H. Eldefrawy, Y. Gadallah and M. Gidlund, "A machine-learning-based technique for false data injection attacks detection in industrial IoT," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 8462–8471, 2020.

[6] V. Borgiani, P. Moratori, J. F. Kazienko, E. R. R. Tubino and S. E. Quincozes, "Toward a distributed approach for detection and mitigation of denial-of-service attacks within industrial internet of things," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4569–4578, 2021.

[7] S. Latif, Z. Zou, Z. Idrees and J. Ahmad, "A novel attack detection scheme for the industrial internet of things using a lightweight random neural network," *IEEE Access*, vol. 8, pp. 89337–89350, 2020.

[8] J. Vávra, M. Hromada, L. Lukáš and J. Dworzecki, "Adaptive anomaly detection system based on machine learning algorithms in an industrial control environment," *International Journal of Critical Infrastructure Protection*, vol. 34, pp. 100446, 2021.

[9] X. Liu, W. Yu, F. Liang, D. Griffith and N. Golmie, "On deep reinforcement learning security for industrial internet of things," *Computer Communications*, vol. 168, no. 5, pp. 20–32, 2021.

[10] A. Alsaedi, N. Moustafa, Z. Tari, A. Mahmood and A. Anwar, "TON_IoT telemetry dataset: A new generation dataset of iot and iiot for data-driven intrusion detection systems," *IEEE Access*, vol. 8, pp. 165130–165150, 2020.

[11] X. Li, M. Xu, P. Vijayakumar, N. Kumar and X. Liu, "Detection of low-frequency and multi-stage attacks in industrial internet of things," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 8820–8831, 2020.

[12] B. Genge, P. Haller and C. Enăchescu, "Anomaly detection in aging industrial internet of things," *IEEE Access*, vol. 7, pp. 74217–74230, 2019.

[13] M. A. Basset, V. Chang, H. Hawash, R. K. Chakrabortty and M. Ryan, "Deep-IFS: Intrusion detection approach for industrial internet of things traffic in fog environment," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7704–7715, 2021.

[14] W. Aoudi and M. Almgren, "A scalable specification-agnostic multi-sensor anomaly detection system for IIoT environments," *International Journal of Critical Infrastructure Protection*, vol. 30, no. 1, pp. 100377, 2020.

[15] T. Vaiyapuri, Z. Sbai, H. Alaskar and N. A. Alaseem, "Deep learning approaches for intrusion detection in iiot networks–opportunities and future directions," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 4, pp. 86–92, 2021.

[16] H. Naeem, F. Ullah, M. R. Naeem, S. Khalid, D. Vasan *et al.,* "Malware detection in industrial internet of things based on hybrid image visualization and deep learning model," *Ad Hoc Networks*, vol. 105, no. 1, pp. 102154, 2020.

[17] Y. Liu, S. Garg, J. Nie, Y. Zhang, Z. Xiong *et al.,* "Deep anomaly detection for time-series data in industrial iot: A communication-efficient on-device federated learning approach," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6348–6358, 2021.

[18] K. Lin, X. Xu and H. Gao, "TSCRNN: A novel classification scheme of encrypted traffic based on flow spatiotemporal features for efficient management of IIoT," *Computer Networks*, vol. 190, no. 5, pp. 107974, 2021.

[19] S. Pundir, M. S. Obaidat, M. Wazid, A. K. Das, D. P. Singh *et al.,* "MADP-IIME: Malware attack detection protocol in IoT-enabled industrial multimedia environment using machine learning approach," *Multimedia Systems*, vol. 16, no. 4, pp. 1, 2021.

[20] G. George and S. M. Thampi, "A graph-based security framework for securing industrial iot networks from vulnerability exploitations," *IEEE Access*, vol. 6, pp. 43586–43601, 2018.