International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2023/223317 Volume 10, Issue 4, July – August (2023)

RESEARCH ARTICLE

IoBTSec-RPL: A Novel RPL Attack Detecting Mechanism Using Hybrid Deep Learning Over Battlefield IoT Environment

K. Kowsalyadevi
Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India.
kowsalyamphilcs@gmail.com

N.V.Balaji
Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India.
fashdean@kahedu.edu.in

Received: 19 June 2023 / Revised: 03 August 2023 / Accepted: 17 August 2023 / Published: 31 August 2023

Abstract – The emerging digital world has recently utilized the massive power of the emerging Internet of Things (IoT) technology that fuels the growth of many intelligent applications. The Internet of Battlefield Things (IoBT) greatly enables critical information dissemination and efficient war strategy planning with situational awareness. The lightweight Routing Protocol for Low-Power and Lossy Networks (RPL) is critical for successful IoT application deployment. RPL has low-security features that are insufficient to protect the IoBT environment due to device heterogeneity and open wireless device-to-device communication. Hence, it is crucial to provide strong security to RPL-IoBT against multiple attacks and enhance its performance. This work proposes IoBTSec-RPL, a hybrid Deep Learning (DL)-based multi-attack detection model, to overcome the attacks. The proposed IoBTSec-RPL learns prominent routing attacks and efficiently classifies the attackers. It includes four steps: data collection and preprocessing, feature selection, data augmentation, and attack detection and classification. Initially, the proposed model employs min-max normalization and missing value imputation to preprocess network packets. Secondly, the enhanced pelican optimization algorithm selects the most suitable features for attack detection through an efficient ranking method. Thirdly, data augmentation utilizes an auxiliary classifier gated adversarial network to alleviate the class imbalance concerns over the multiple attack classes. Finally, the proposed approach successfully detects and classifies the attacks using a hybrid DL model that combines LongShort-Term Memory (LSTM) and Deep Belief Network (DBN). The performance results reveal that the IoBTSec-RPL accurately recognizes the multiple RPL attacks in IoT and accomplished 98.93% recall. It also achieved improved accuracy of 2.16%, 5.73%, and 6.06% than the LGBM, LSTM, and DBN for 200K traffic samples.

Index Terms – Internet of Battlefield Things (IoBT), RPL, Multiple Routing Attacks Detection, EPOA, Hybrid Deep Learning, Attack Detection and Classification.

1. INTRODUCTION

The computing capability of the Internet of Things (IoT) enables countless next-generation applications and improves the human lifestyle by connecting physical objects with the virtual world through the Internet [1]. The emerging IoT is growing exponentially to realize many real-time smart applications. Among them, defense and public safety with IoT receive high attention in which the IoT transforms the fundamental military surveillance system using underlying intelligent wireless sensing. The Internet of Battlefield Things (IoBT) is a new breed of extension that introduces IoT intelligence in military surveillance security and situational awareness [2].

The Low Power Lossy Networks (LLNs) receive significant attention in IoT, as it provides scalable and ubiquitous communication infrastructure [3] [4]. The Routing Protocol for LLNs (RPL) is highly suitable for IoTs in which reliable routes are constructed among the low-powered IoT devices using objective functions and Destination-Oriented Directed Acyclic Graphs (DODAGs) [5].

Despite the potential advantages of the RPL routing mechanism, limited security opens the door to vulnerabilities. The RPL is designed with a lack of specific security features, and thus, it increases the security risks in IoBT-RPL applications. Also, the limited computing capability of IoT devices leads to various security attacks. The attackers intend to decrease RPL protocol performance and increase the risks of secure RPL deployment. Hence, detecting and classifying the attacks in RPL-based IoT is essential to improve performance efficiency.

Enabling the routing operation in the IoBT-RPL-based network is a significantly challenging task, primarily the resource constraints of the heterogeneously interconnected devices in the presence of adversaries. Many RPL attack detection methods are introduced to detect and mitigate malicious behaviors [6]. Several detection mechanisms have been proposed, such as acknowledgment-based, trust-based, and many more. Among them, intrusion detection methods have received high attention [7] [8].

Enabling artificial intelligence to combat attacks has recently become a prominent research area. Due to the detection accuracy, intrusion detection with Deep Learning (DL) has been a highly promising solution for RPL-IoT attack detection [9]. The DL algorithms can handle massive heterogeneous IoT data and have the capability of detecting complex attacks with significant attack detection efficiency.

Distributed sharing and processing of the workloads across edge nodes may enhance detection in terms of quality, including detection speed, accuracy, and response time. The computation of learning steps for detection and prevention requires sophisticated high-end computation of DL networks performed in resource-constrained devices [10].

1.1. Contributions

The summary of contributions is discussed as follows:

- The primary objective of the proposed IoBTSec-RPL is to detect and classify multiple RPL attacks over IoBT scenarios and to enhance the performance of military operations. To accomplish this goal, the proposed model designs the attack classification system with the enhanced feature selection and hybrid deep learning model.
- IoBT is an emerging technology with no recent and specific IoBT datasets. To deal with the unavailability of IoBT-specific datasets, the proposed model generates a novel dataset using Cooja Simulator.
- The proposed model applies the Enhanced Pelican Optimization Algorithm (EPOA) technique to select the most relevant attack detection features, ensuring minimal computational burden and accurate classification. Also, it handles the class imbalance through ACGAN based data augmentation process for the selected features.
- To detect and classify multiple RPL attacks successfully, the proposed model combines the LSTM and DBN deep learning algorithms. The hybrid model increases the attack detection accuracy by generalizing the classification even when there are unknown attacks, then fine-tuning the classification of multiple attacks.
- Finally, the simulation results are validated for the proposed IoBTSec-RPL using accuracy, precision, recall, Matthews correlation coefficient, and cross-entropy.

1.2. Paper Organization

The structure of the paper is presented as follows. Section 2 surveys the secure routing works proposed for the RPL-IoT environment. Section 3 provides the system and attack model of IoBTSec-RPL. Section 4 overviews the proposed model and explains the proposed mechanisms in detail. Section 5 validates the proposed model, and the paper's conclusion with possible future research directions is presented in section 6.

2. RELATED WORK

The attacks on the RPL routing and its mitigation methods are specially covered in this section and provide detailed research progress in different machine learning and DL-based attack detection methods.

2.1. Machine Learning-based Defence Mechanisms in Handling RPL Attacks

The research work [11] proposed a lightweight multiclass classification using a gradient-boosting model for RPL-specific wormhole and rank attack discovery. This model developed a unique dataset to model the attacks, utilizes optimum feature selection, which enhances the performance of the model. However, while creating Machine-Learning (ML)-based attack detection methods, the concurrent processing of both attacks is not considered. An improved Protocol-specific and SN-inherited Attack Detection (ProSenAD) [12] model is proposed to overcome the identified gap in the previous work. It mainly enhances the ML models to concurrently detect protocol-specific and senor network inherited attacks using multiclass classification. In [13], an intrusion detection system is proposed using an extreme parameter-optimized Light Gradient Boosting Machine (LGBM) for version attack classification. The proposed model created a huge VNA dataset, an extracting features technique, and maximum parameter optimization to enhance the VNA detection accuracy. The suggested LGBM model has less training and testing time and is appropriate for resource constraint IoT devices.

The work in [14] presents an efficient machine learning-based IDS for IoT. It monitors the network activities to detect the attacks like decreased rank, blackhole, sinkhole, version number, and selective forwarding. The machine learning-based multi-attack detection strategy selects the most relevant features based on genetic recursive and classifies the attacks using the fuzzy k-nearest classifier. Thus, it improves the attack detection accuracy with less false positive rate. The work in [15] introduces an Ensemble Learning based Network Intrusion Detection System named ELNIDS architecture to determine the routing attacks against IPv6-based RPL networks. It uses four ML classifiers to design the ensemble classification model. The ELNIDS exploits the RPL-NIDDS17 dataset that comprises the attack packet traces of the following attacks: Sinkhole, Clone ID, Local Repair, Selective Forwarding, Blackhole, Sybil, and Hello Flooding, to its evaluation.

Consequently, the work in [16] proposes an XGBoost ML algorithm-based IDS model to detect the version number, hello flood, and decreased rank attacks. The XGBoost-based IDS models detect attacks more accurately than other ML-based IDS models. The work in [17] designs an ML-based anomaly IDS model to determine the most critical RPL-IoT attacks. It generates an IoT-based dataset using the Cooja simulator to evaluate and fix minor attack classes to improve accuracy.

2.2. Deep Learning-based RPL Attack Detection Approaches

The work in [18] presents a Gated Recurrent Unit (GRU) based deep learning strategy to detect the hello flooding attacks against the RPL network. Compared to other deep learning models, the GRU has a simple structure and allows appending new gates with light codes, resulting in high learning accuracy with minimum time. In [19], a robust RPL-IoT attack detection framework has been presented aiming to detect identity impersonation attacks. The classical RPL-IoT is vulnerable to misidentifying. Therefore, the robust model employs unsupervised pre-training strategies to choose key characteristics from the network samples RPL-IoT.

Further, it utilizes a Dense Neural Network (DNN) deep learning model to enhance malicious attack detection and classification accuracy. Consequently, a deep learning IDS-based RPL rank attack detection mechanism has been introduced in [20]. It utilizes the Cooja simulator to generate an IoT dataset with normal and abnormal traffic. Further, it analyses the network traffic using Multi-Layer Perceptron (MLP) and successfully classifies the rank attackers. The work in [21] introduces a novel deep learning-based early-stage detection (DL-ESD) to detect multiple attacks like rank, version number, and hello flood attacks over RPL-IoT. It generates a novel IoT routing attack dataset named IRAD for efficient detection. Further, it uses linear discriminant analysis (LDA) to extract the most relevant features from IRAD. It trains the *K*-nearest neighbors (KNN), logistic regression (LR), naïve Bayes (NB), support vector machine (SVM), and MLP classifiers using the extracted IRAD training samples. In addition, it detects multiple attacks and classifies them under particular categories using deep learning algorithms.

A CNN-based anomaly IDS model has been proposed in [22]. It enhances the security performance of IoT networks by training and testing the CNN using two different datasets: NID and BoT-IoT. The work in [23] proposes an unsupervised ensemble Deep Belief Network (DBN) learning strategy to determine the IoT attacks. It detects the attacks over the unlabelled dataset and performs the learning process with the system-generated labeled dataset. Also, it employs a feature reduction mechanism to decide the most relevant features from the system-generated dataset and improves the learning accuracy. A reliable DL-based routing attack detection model for industrial RPL-IoT has been presented in [24]. It includes a two-stage combination model named Generative Adversarial Network-Classifier (GAN-C) for attack detection. The two-stage model is designed using GAN and SVM algorithms. It includes a parallel learning and detection method that effectively supports the deep learning model over resource-limited IoT devices. The GAN-C attack detection model detects the attack events through real-time monitoring. Consequently, the work in [25] utilizes an auxiliary classifier GAN (ACGAN), for data augmentation and determines the machine faults. ACGAN has a simple training structure and can potentially neglect the gradient vanishing issue during training. Thus, it improves the training accuracy with auxiliary data and maximizes detection accuracy.

2.3. Problem Statement

IoBT is critical in improving the effective dissemination of military information among the heterogeneous network devices on battlefields. Designing a routing protocol in the IoBT-RPL-based network is significantly challenging, notably the resource constraint heterogeneous connectivity of devices in the presence of adversaries. The efficiency and attack detection in dynamic war scenarios relies on real-time data collection in the vulnerable military network. The information observed by various devices like sensors and wearable devices of soldiers is transferred through the device to device communication and edge intelligence. The power of deep neural network architecture is utilized efficiently to detect attackers. The DL algorithms can handle massive heterogeneous IoT data and have the capability of detecting complex attacks with significant attack detection efficiency.

3. SYSTEM AND ATTACK MODEL

The RPL-IoBT network is modeled as a communication graph G (N, E). The term N denotes the number of IoBT devices, and E refers to the RPL communication links among N. Let us consider that the proposed IoBT scenario comprises N numbers heterogeneous devices. Each IoBT device corresponds to various battlefield things like surveillance cameras (Dc) and soldier-wearable intelligent devices (Ds-w), where N€{Dc, Ds-w}. The devices are static and dynamic. The devices Dc are static, and the devices Ds-w are dynamic. Thus, it creates frequent topological changes, and the RPL can effectively handle such issues. The different IoBT devices are primarily characterized by the usage of transmission power or communication range. The IoBT devices can establish internet-based communication with the assistance of the ISR satellite, which rounds the earth and provides a wide range of internet coverage. The communication range of sensor and soldier wearable devices is C, and each IoBT device has adequate power to establish device-to-device communication.

Since the IoBT devices are equipped with meta material wearable patch antenna to charge their batteries. The devices within the communication range can establish direction communication links, and the devices out of the communication range need routers or intermediates. The IoT device has a communication range denoted as $r_{N,}$ and the m number of devices is distributed per kilometer. The proposed model collects the network information of various IoBT devices using cooja-based simulation to construct the novel dataset over a certain period. A typical IoBT architecture is shown in Figure 1.

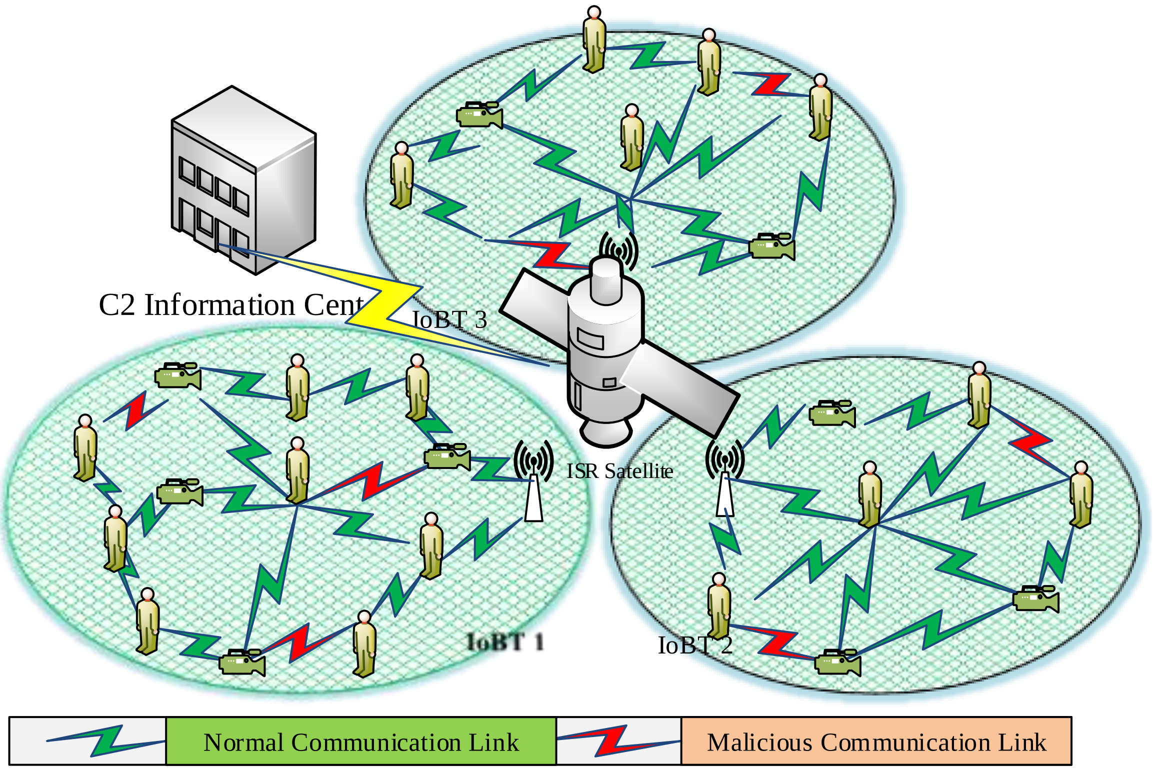


Figure 1 IoBT Architecture

The collected dataset comprises both normal and attack samples. The attack samples include the following attacks.

- Rank Attack (DR): In this type, the malicious nodes attempt to advertise minimum or maximum rank values to prove it is proximity to the root node. Thus, it motivates the legitimate nodes to select the malicious nodes as their parents during RPL construction.
- Hello Attack (HF): To make themselves known to all nearby nodes, hostile nodes disseminate hello packets by leaking DIO packets into the network. By providing Hello packets with better network statistics, the attacker deceives other nodes in the network.
- Sink Hole (SH): In this attack, the adversary attempts to connect to another node by announcing an optimal rank, allowing them to build a fake channel that draws traffic via it successfully.
- Black Hole Attack (BH): This attack also causes many packets to be discarded when routed via the malicious node, impeding network data flow.
- Version Number Modification (VN): To purposefully disrupt the current DODAG topology, the attacker modifies the DODAG's version number.

4. PROPOSED METHODOLOGY

The proposed IoBTSec-RPL intends to detect multiple RPL attacks using a hybrid deep learning model and improve military surveillance performance. As shown in Figure 2, the proposed model includes data collection and preprocessing, feature selection, data augmentation, and attack detection and classification. Initially, the proposed model collects data from the RPL-IoBT network and preprocesses the dataset using EPOA. The EPOA neglects the irrelevant features from the constructed dataset for attack detection and mitigates the computation time without degrading the classification accuracy. Subsequently, it performs the data augmentation using ACGAN, in which high-quality class-specific network traffic is generated. Further, the augmented dataset is provided as input to the hybrid learning model for recognizing abnormal network traffic and classification of multiple RPL-IoBT attacks by utilizing the advantages of both the LSTM and DBN models.

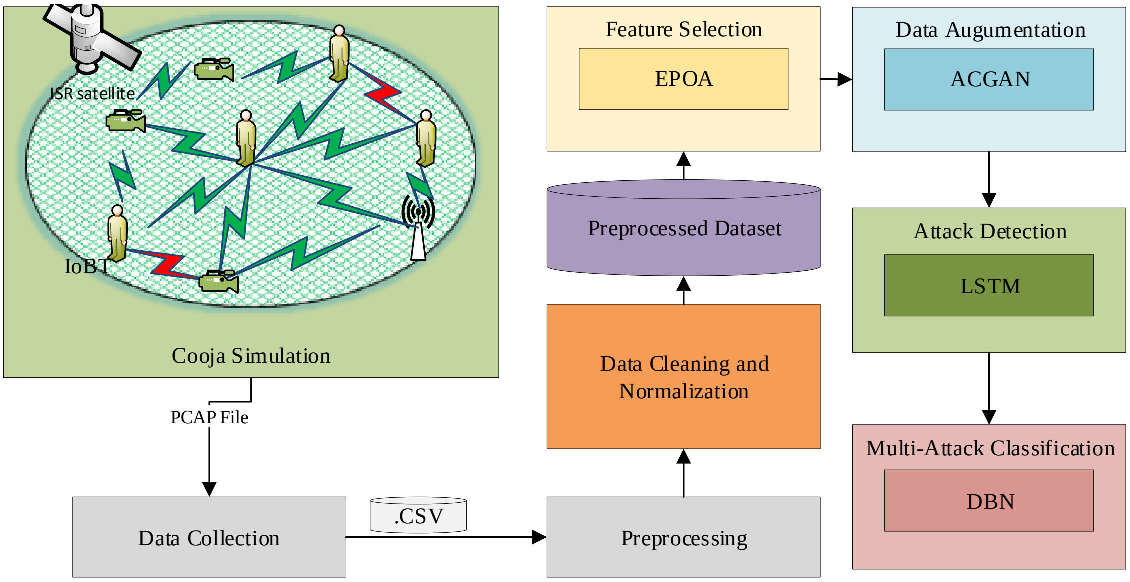


Figure 2 Proposed Methodology

4.1. Data Collection and Preprocessing

An ordinary RPL network with an edge node is first set up without any malicious node. One hundred separate pcaps are created when the test scenario has been performed one hundred times. It generates the dataset's examples that correspond to typical traffic flow. The pcap file generates the dataset containing the data passing over the network. Many simulations were conducted for each situation ranging from 10 to 15 times. One row of the dataset is constructed using the Pcap captures of each scenario. The label is the class value used to train the deep learning network and represents a particular attack vector.

Table 1 Attacker Modeling in IoBT Simulation

Attacks	Attacker Nodes
Rank	14, 24
Version	27, 47
Blackhole	7, 30
Hello Flood	42, 52
Sinkhole	17, 37

Table 1 shows the intruder or attacker node IDs modeled in the simulated IoBT environment. The gathered information is analyzed and stored in Comma-Separated Values (CSV) format.

Preprocessing: Preprocessing mainly focuses on data cleaning, feature generation, and feature selection. Feature reduction decides a subset of relevant characteristics. Generally, the collected dataset comprises various features that do not contribute to attack detection. Providing the collected dataset straightly input to the detection model increases the learning time and decreases the detection accuracy. Data preprocessing is essential to select the features contributing to attack detection.

4.2. Feature Selection Using EPOA

Feature Selection (FS) is one of the essential preprocessing for ML, as it eliminates irrelevant, redundant and noisy features from learning input and increases the model accuracy with high speed. The proposed model utilizes Pelican Optimization Algorithm (POA) to select features. It works based on the hunting behaviors of pelicans. POA's primary advantage is adjusting parameters such as fast convergence speed and simple calculations. The proposed model enhances the original POA [26] and applies it to select the most relevant features to attack detection.

Algorithm 1 explained the proposed Enhanced POA (EPOA). The EPOA utilizes correlative information rather than randomly initializing the pelicans and prey in the search space. Before initializing feature subsets, the correlation is computed between the features and class labels and assigns the positively correlated feature set with labels as the alternative feature subset in the local search space to avoid the scattered exploration area and increase the computation efficiency. As a result, the improvement in the pelican movement selects the optimal set of features in minimal computation time. The dataset has F features; the feature set is defined as k∈ {f1, f2,..., fn}. Thus, the proposed approach effectively reduces the dimensionality of the network traffic with an optimal set of features to facilitate multiclass attack detection in the IoBT environment.

```
Input the optimal problem information
Determine the population size of POA as number of features (F) and number of iterations (T)
Initialize the pelican position and calculate the objective function
For each iteration (t), $t \in T$ do
Generate the pelican position of prey based on positive correlation in between features and labels
If feature subset f2 is positively correlated with label then
Initialize the position of f2 as the alternative feature set for f1
End
For all the number of features (F), $k \in \{f1, f2, ..., fn\}$
Phase 1 : Moving towards prey (exploration phase)
For each feature subset (k), $k \subseteq F$ do
Calculate new position of pelicans with the higher accuracy range and minimal number of features
End
Update the population of pelican or feature (f)
Phase 2 : Winging on the water surface (exploitation phase)
For each feature subset (k), $k \subseteq F$ do
Calculate the position of pelicans correlation information and collective intelligence between new position of pelican search space
End
Update the best candidate solution to move the pelican towards reach of its global best solution
End
Update the best candidate solution based on the fitness function on all iterations
```

Algorithm 1 EPOA Feature Selection Process

4.3. ACGAN-Based Data Augmentation

The training set significantly impacts the DL model performance. A DL architecture has several hidden units and many untrained free parameters. The deep architecture must be trained properly to make correct predictions, and completely retraining a big network often takes a lot of balanced data. In reality, train samples taken from various machine states are often out of balance. It is somewhat challenging to train a deep model to provide an accurate forecast of network conditions when there is limited training data, particularly imbalanced data. GAN solves imbalanced data categorization by producing minor classes. GANs have effectively been used in producing synthetic samples and have shown a strong capacity to produce realistic data. By creating false data comparable to the actual data, the proposed approach augments the reduced data without imbalanced classes, and thus, the design of ACGAN [27] avoids over fitting. Additionally, the generative architecture model aids in understanding the distribution of the original data and offers a fresh viewpoint for identifying attacks on the RPL Protocol for IoT networks. The whole GAN architecture's objective is summarized as discussed in equation (1).

$$\mathrm{Goal} = \arg\min_{\mathrm{G}} \max_{\mathrm{D}} \mathrm{L}(\mathrm{G}, \mathrm{D}) \qquad (1)$$

An appropriate pair of discriminator (D), as well as a generator (G), may be obtained by training ACGAN using Stochastic Gradient Descent (SGD) and updating model parameters depending on the objective function. It utilizes extra class labels for the generator and discriminator, resulting in a modified architecture of the standard GAN. In the ACGAN, the discriminator is additionally coupled with an auxiliary component that results in precise class labels, enabling the enhanced discriminator to distinguish between distinct classes and with the reference of input network traffic. Figure 3 shows the ACGAN data augmentation process.

Auxiliary classifier generative adversarial networks are a variation that blends conditional class architecture with auxiliary networks for classification. ACGANs may produce high-quality data and provide labeled data immediately, unlike ordinary GANs. To create fake samples, the generator statistically utilizes the noise ‘z’ and label ‘l’, and the discriminator generates probabilities based on the information source and corresponding labels. The quality of the produced data samples is important since the purpose of employing the GAN architecture is to provide appealing data samples to perform data augmentation. It is crucial to assess how closely produced data resembles actual data. Nevertheless, it is not appropriate to directly calculate the score using a model that has received training on natural images for series data signals, particularly mechanical sensor signals. However, considering the inherent feature of sensor information, time domain features and relative frequency can reflect frequency components to some extent. As a result, these statistical traits were used to develop the assessment measures. To produce high-quality fake sensed network traffic for data augmentation, create a framework utilizing ACGAN. The created data may then supplement imbalanced datasets for further applications. The whole dataset is constrained by the data from the minor class in practice, making an imbalanced training dataset unsuitable for training models that perform fault classification. As a result, the original dataset may be enriched with high-quality data produced by GAN-based models, which helps with defect identification. This essay focuses on the production of synthetic data and sample assessment. In the ACGAN, the combination of the generator and discriminator enforces the generator to produce samples from latent space with a particular label. Before sending the generated samples to the discriminator, they are combined with actual training data. Discriminator can produce two labels for each sample, one showing whether the sample is genuine (output 1) or fake (output 0) and the second denoting a certain class. As a result, the proposed approach augments the input network traffic from the perspective of class-specific attack patterns in the IoBT platform.

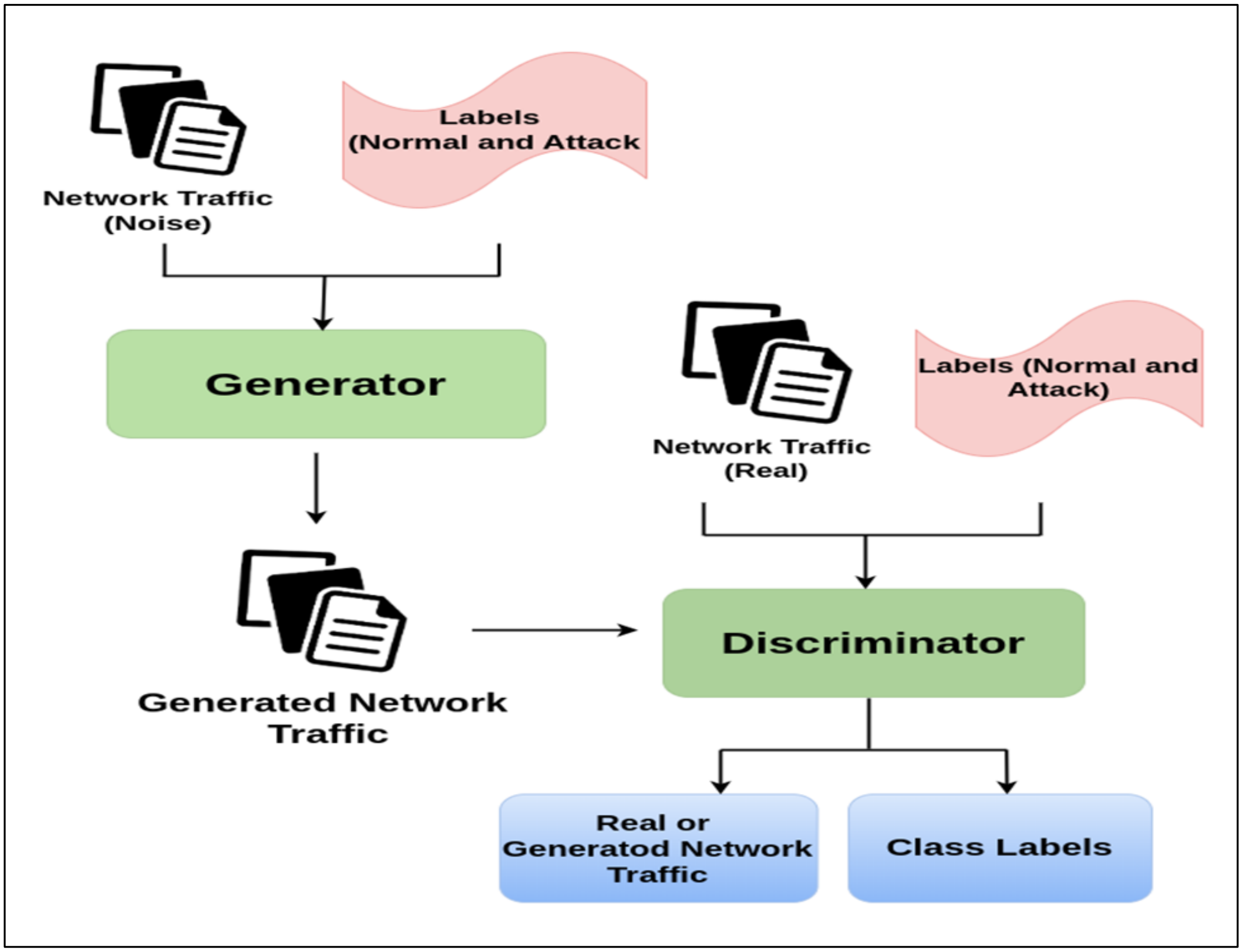

Network Traffic (Noise)
Labels (Normal and Attack
Generator
Network Traffic (Real)
Labels (Normal and Attack)
Generated Network Traffic
Discriminator
Real or Generatod Network Traffic
Class Labels

Figure 3 ACGAN for Data Augmentation

4.4. Attack Detection and Classification

The hybrid DL model used in the proposed model combines the advantages of two different DL models of LSTM and DBN for accurate multiclass attack detection. The proposed IoBTSec-RPL sequentially integrates the LSTM and DBN to detect the attacks under multiple classes. Firstly, adopting the LSTM model is responsible for verifying whether the augmented data traffic is normal or abnormal by comparing it with its binary-labeled training data. Consequently, the proposed IoBTSec-RPL can recognize normal and abnormal network traffic, including known and unknown malicious patterns. Secondly, the DBN classifies the attacks under various attack categories based on the output provided by LSTM. Hence, the known and novel attacks are also identified and classified under the corresponding label. Figure 4 illustrates the proposed hybrid DL model.

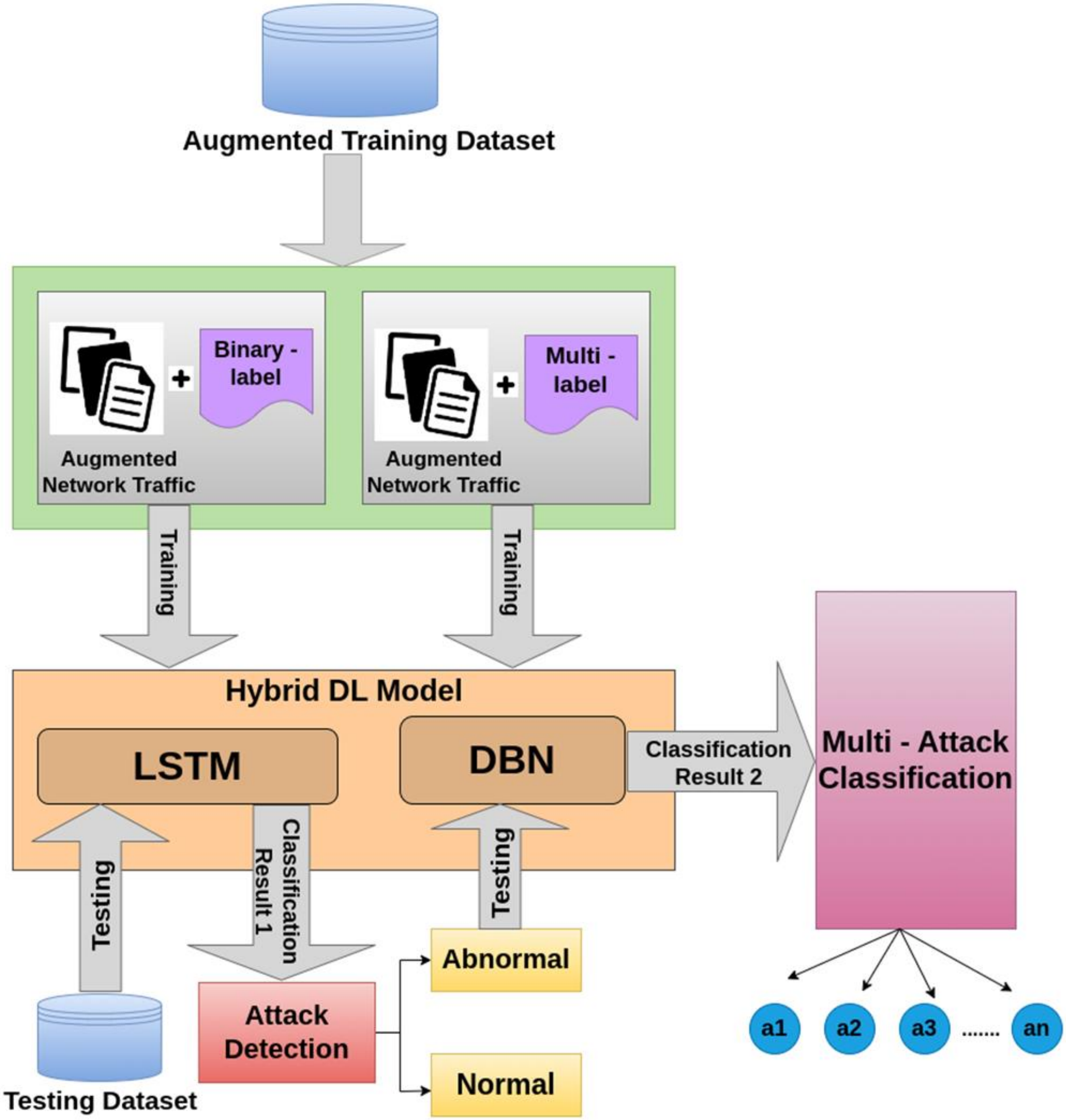


Figure 4 Proposed Hybrid DL Model

LSTM – Attack Detection: The proposed model utilizes the LSTM model to detect malicious behaviors in the RPL-IoBT environment. The proposed approach can recognize all the unknown attack patterns by providing the binary labels as the normal and abnormal network traffic patterns to the LSTM model. It is accomplished by learning possible normal behaviors instead of training the model with known attack behaviors. The LSTM network has feedback connections that are more advantageous than fundamental neural networks. The LSTM comprises multiple layers: input, hidden, and output. It exploits the series of gates responsible for controlling the input data sequences provided by the training dataset. The LSTM includes three types of gates: input, forget, and output. The forget gate maintains and decides the cell state bits of long-term and short-term memory, which benefit the previous output and hidden layer. The second one is the input gate that selects the input for the current LSTM, and the third is the output gate that generates the corresponding output. The proposed IoBTSec-RPL gives the unlabeled training samples as input to the LSTM, and the LSTM decides the forget gate value based on the input sequence and previous state output. Therefore, the LSTM IoBT node is fed with the output of the previous hidden state (h (t-1)), current input x (t), and cell state of the previous state (c (t-1).

The cell state comprises useful hidden information LSTM. Consequently, the data is passed from three gates in LSTM: input, forget, and output. With the assistance of the three LSTM gates, it maintains potential information or forgets irrelevant information about the old cell state and estimates the next output. The output of LSTM lies under two categories that are normal and abnormal. Algorithm 2 explains the LSTM attack detection process. Further, the LSTM provides abnormal results in recognizing new attack classes during the classification of multiple attack patterns by the DBN model. Algorithm 2 illustrates LSTM Attack Detection.

```
Input: Binary labeled Training set with Normal and Attack Classes, x (LSTM);
Output: Normal and Abnormal traffic detection
Function LSTM-Training;
Train the LSTM with binary labeled input data;
Fixes the bias and error value using forget, input, and output gates;
Function LSTM-Testing;
Provide testing set to LSTM;
For (compare each test input with bias and error value) {
Detect the abnormal traffic;
};
```

Algorithm 2 LSTM Attack Detection

DBN - Multi-attack Classification: After discriminating the attack behaviors from the normal patterns by the LSTM, the proposed model exploits the DBN to classify the attacks under multiple categories, which are rank, version, blackhole, hello flood, and sinkhole. In the DBN, the Restricted Boltzmann Machine (RBM), a two-layer network, becomes the initialization unit. Further, the visible and invisible states of DBN are formed. The RBM provides the relations between layers. The proposed DBN consists of two steps: labeled sample-based training and multi-attack classification. The augmented multiclass training samples are provided as input to DBN during the DBN training. During attack detection, the abnormal sample output of LSTM is provided as the test input to the DBN, and the DBN classifies the attacks according to the training knowledge. The multi-attack classification using DBN is explained in Algorithm 3.

```
Input: Multi-labeled Training Set x(DBN), Abnormal traffic output of LSTM y(LSTM(a)), and Test Set y(DBN);
Output: Multi-Attack Classification
The IoBT node do {
Initialize DBN for attack Classification;
Function DBN-Training;
Train the DBN using multiple labels;
Train the DBN with attack thresholds ($Th_{R,}$ $Th_{HF,}$ $Th_{Bl}$, $Th_{Sl}$,and$Th_{V,}$);
If (x(DBN==0)) {
Data is normal;
Else if (Rank Value>$Th_R$) {
Rank attack;
Else if (HF Count>$Th_{HF}$OR(HF Count<$Th_{HF}$) {
Hello Flood Attack;
Else if (Count>$Th_{Bl}$and Count>$Th_{Bl}$) {
Black hole attack;
Else if (Count>$Th_{Sl}$and Count>$Th_{Sl}$) {
Sink hole attack;
Else if (Version Number≠ $Th_V$) {
Version Number Attack;
Else if (New behaviors observed) {
Novel attack and new labeling;
}}}}}}};
Function DBN-Testing:
Provide y(DBN) as input to DBN testing process;
Compare the testing data with training data;
If (y(DBN)=0) {
        Normal behavior;
Else {
```

```
Classify the attacks under different categories a1, a2, a3, a4, and an;
}};
```

Algorithm 3 DBN Multi-Attack Classification

5. RESULTS AND DISCUSSION

The effectiveness of the proposed IoBTSec-RPL model is evaluated by generating the data using contiki/cooja-based simulations. Python libraries are used to build machine-learning models. The IoBTSec-RPL was implemented using Ubuntu 18.04 LTS operating system with the Intel i3 2.5GHZ CPU and 8 GB memory. The IoBTSec-RPL is mainly proposed for the IoBT environment in which intelligent communication devices are used, and there are no publicly available datasets suitable for the IoBT environment. Therefore, the proposed model constructs a novel dataset by collecting the data from 60 nodes in the IoBT environment. The nodes are soldiers, surveillance cameras, and armor vehicles, and they are static and dynamic locations. The proposed model exploits the random waypoint mobility model to obtain the dynamic locations of moving IoBT nodes like soldiers and armor vehicles. The proposed model gathers the network information from IoBT to construct the dataset. The collected dataset should comprise both normal and attack traffic. Therefore, the IoBTSec-RPL considers a particular number of IoBT devices and 20% of total node density as attackers to malicious scenario creation. To evaluate the performance of IoBTSec-RPL accurately, the proposed model creates different scenarios. For an accurate evaluation, the proposed model is compared with the widely used LGBM model for attack detection, LSTM, and DBN. Moreover, the simulation network parameters are illustrated in Table 2.

Table 2 Simulation Network Parameters

Parameters	Values
Simulator (Version)	Instant Contiki 3.0 / Cooja
Coverage Area	200m * 200m
Number of Nodes	60
Malicious Nodes	20% of total node density
Communication range of IoBT devices	50m
Communication range of sink	50m
Nodes Deployment	Random
Mobility Model	RandomWayPoint
Delay Starting	Less than 2s
Simulation Time	180s
Network Protocol	IPv6 based
Routing Protocol	RPL

5.1. Performance Metrics

The effectiveness of the proposed model is estimated in terms of Classification Accuracy (A), Precision, Recall, and Matthews Correlation Coefficient (MCC), as discussed in Equations (2-5).

Accuracy: It is the percentage of correctly identified attackers to the total number of attackers. Equation (2) provides a formula for calculating accuracy, where TP = True Positive, FP = False Positive, TN = True Negative, and FN = False Negative.

$$A= \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

Precision: It is the measure of correctly classified attackers to totally classified attackers. It is measured using the following formula.

$$P = \frac{TP}{TP+FP} \tag{3}$$

Recall: It measures correctly identified attackers to the actual number of attackers. The sensitivity/recall is calculated using equation 4 that follows.

$$R = \frac{TP}{TP+FN} \tag{4}$$

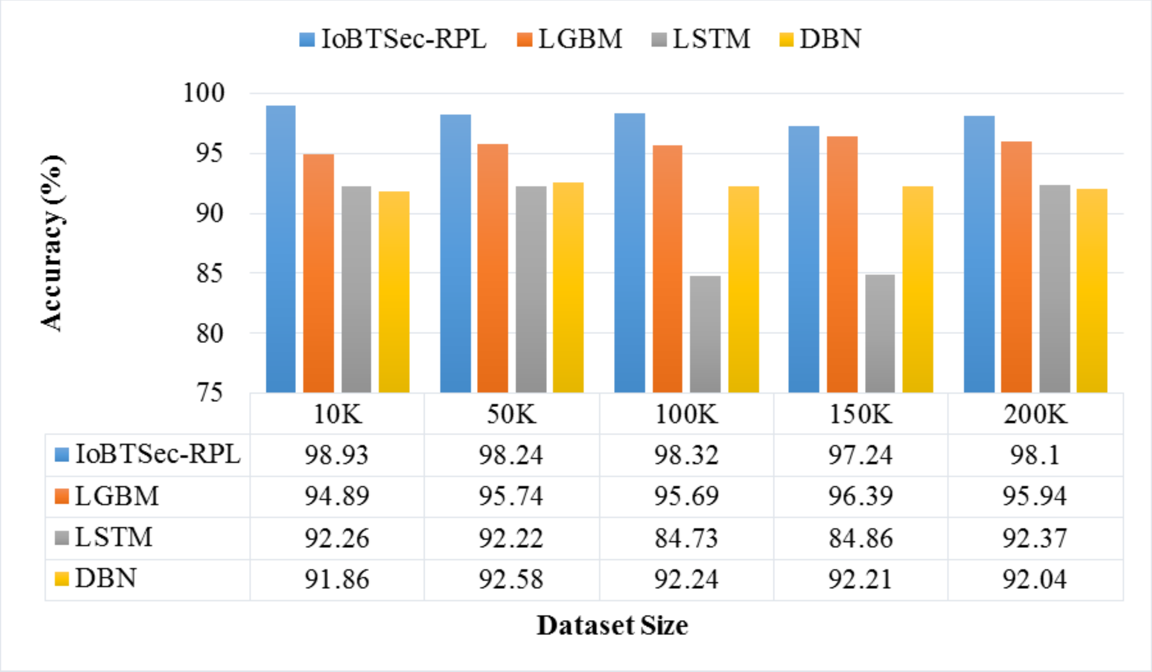
MCC: It estimates the correlation coefficient between the observed and detected classifications.

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5}$$

5.2. Simulation Results

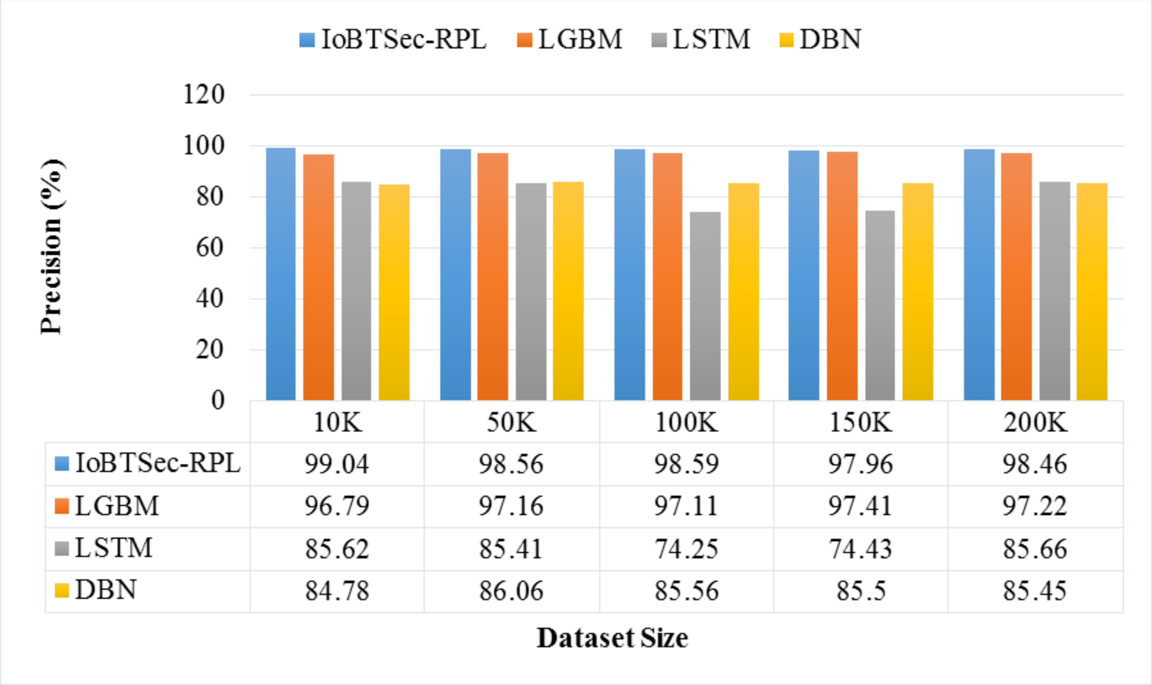
Figure 5 shows the accuracy comparison results of the proposed IoBTSec-RPL, LGBM, LSTM, and DBN obtained by varying the dataset size from low to high. The comparative models rapidly increase and decrease the accuracy by adjusting the dataset size from 10000 to 200000. In contrast, the IoBTSec-RPL accomplishes 98.93% and 98.1% for 10K and 200Ksamples in the dataset, respectively, illustrating that the proposed approach optimally maintains the attack detection performance with marginal differences in the accuracy even when the dataset size increased from 10K to 200K samples. The main reason is that the proposed work utilizes the hybrid classifier for attack detection. The advantages of LSTM and DBN models improved the attack classification accuracy from the reduced network features.

Unlike IoBTSec-RPL, the existing models are single-learning models. There is a possibility of higher misclassification due to a lack of recognizing the generalized attack behaviors from the normal network traffic. For instance, the IoBTSec-RPL improves the 2.16%, 5.73%, and 6.06% than the LGBM, LSTM, and DBN, respectively, for 200K traffic samples in the entire IoBT network traffic dataset.



	10K	50K	100K	150K	200K
🟦 IoBTSec-RPL	98.93	98.24	98.32	97.24	98.1
🟧 LGBM	94.89	95.74	95.69	96.39	95.94
■ LSTM	92.26	92.22	84.73	84.86	92.37
🟨 DBN	91.86	92.58	92.24	92.21	92.04

Figure 5 Dataset Size Vs. Accuracy



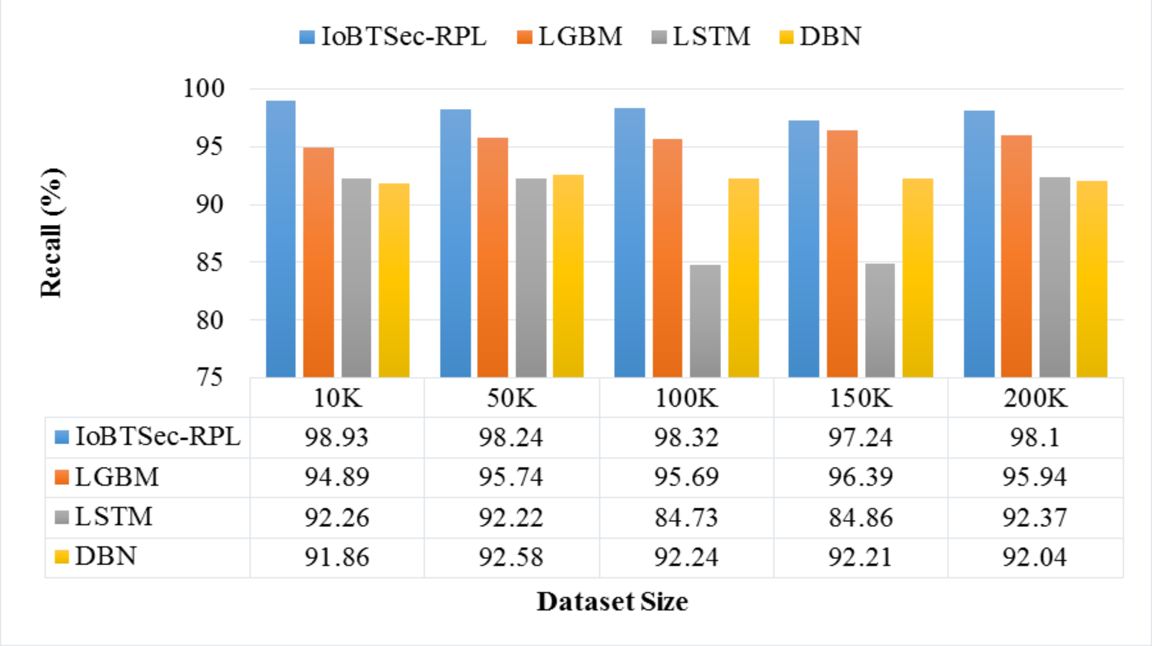
	10K	50K	100K	150K	200K
🟦 IoBTSec-RPL	99.04	98.56	98.59	97.96	98.46
🟧 LGBM	96.79	97.16	97.11	97.41	97.22
■ LSTM	85.62	85.41	74.25	74.43	85.66
🟨 DBN	84.78	86.06	85.56	85.5	85.45

Figure 6 Dataset Size Vs Precision

Figure 6 illustrates the comparison results of precision of proposed IoBTSec-RPL, LGBM, LSTM, and DBN obtained by varying the dataset size. The precision values of LGBM, LSTM, and DBN obtained comparatively minimal precision values than the proposed IoBTSec-RPL while varying the dataset size from 10K to 200K. Even though the dataset size is increased to 200K, the IoBTSec-RPL attains 98.46% precision by optimally reducing the network features and then recognizing the multiple attacks from only the abnormal samples recognized by the LSTM. As a result, the precision of the proposed IoBTSec-RPL is higher at 2.25% compared with the LGBM model when the dataset size is 10000. By designing the EPOA and ACGAN for feature selection and data augmentation in the proposed approach, the precision value becomes optimal for different ranges of the input dataset. For example, the IoBTSec-RPL increases the precision by 24.34% and 13.03% more than the independent LSTM and DBN models, respectively, when the dataset size is 100K.

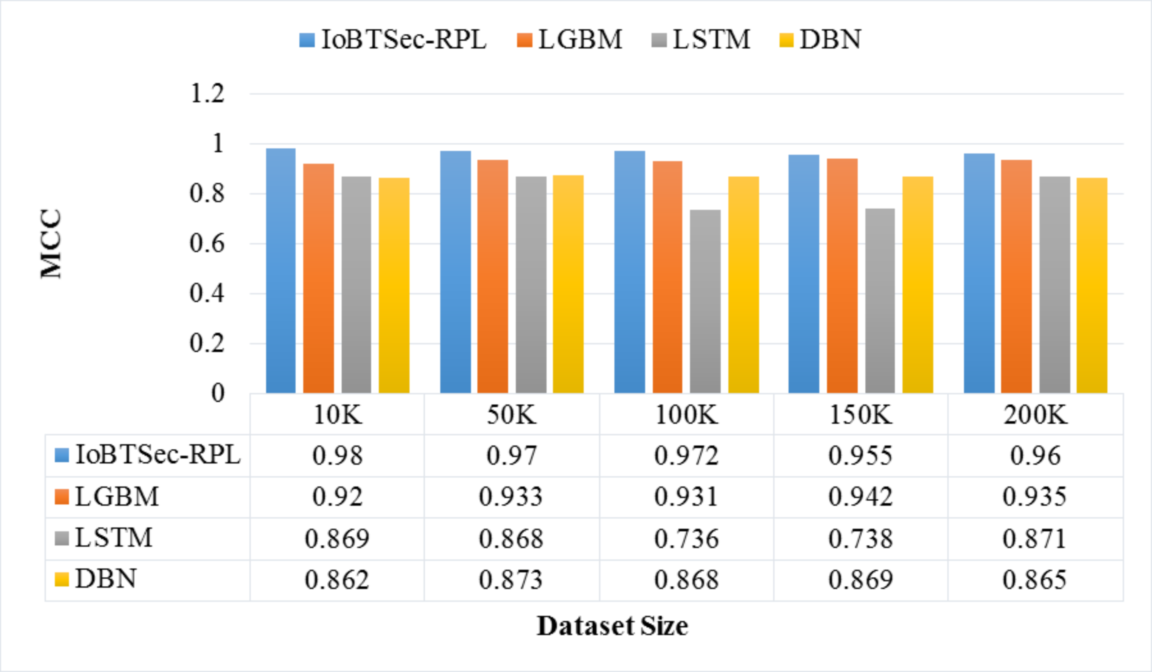
Figure 7 shows the recall comparison results of IoBTSec-RPL, LGBM, LSTM, and DBN models. The proposed IoBTSec-RPL achieves 98.93% recall value for 10K and 97.24% recall for 150K dataset size with marginal difference in its recall performance. In contrast, LSTM abruptly degrades its recall value from 92.26% to 84.36% when increasing the dataset size from 10K to 150K. It is because the proposed approach optimally selects the features and augments the network traffic with the consideration of the attack labels, facilitating the maintenance of potential attack patterns in the training knowledge without burdening the learning model. Moreover, the hybrid learning model-based attack detection and classification minimizes the error rates and improves 12.38% higher recall for hybrid LSTM-DBN compared to the LSTM model when the dataset size is 150K. Moreover, the IoBTSec-RPL improves the recall by 2.5%, 6.02%, and 5.66% compared to the LGBM, LSTM, and DBN for the 50K dataset size scenario, proving that the proposed approach gradually reaches its optimal value regardless of the dataset size compared to the baseline models.

Figure 8 depicts the MCC comparison results of proposed and comparative baseline models obtained by varying the dataset size from 10K to 200K. The MCC of the proposed IoBTSec-RPL yielded 96.74% of average MCC for a different number of samples in the dataset, whereas the LGBM obtained 93.22% of average MCC even though the LGBM model gradually increased its performance with the increase of the dataset size. The high number of samples in the dataset does not contribute to the improvement in the classification accuracy. Hence, the proposed approach accomplishes higher MCC and maintains the optimal value even when there is a minimal set of samples or a huge number of samples. Furthermore, the comparative LGBM model obtains a higher MCC of 0.942 when there are 150K samples than the 0.931 MCC obtained in 100K samples, showing that comparatively better performance than the proposed approach but suffers with the local convergence that proves in the performance degradation of MCC value when the dataset size is 200K. By the design of the EPOA, the proposed IoBTSec-RPL avoids premature convergence and improves the multiclass attack detection accuracy. Thus, the MCC of the proposed IoBTSec-RPL achieves better performance than the existing LGBM classifier.



	10K	50K	100K	150K	200K
IoBTSec-RPL	98.93	98.24	98.32	97.24	98.1
LGBM	94.89	95.74	95.69	96.39	95.94
LSTM	92.26	92.22	84.73	84.86	92.37
DBN	91.86	92.58	92.24	92.21	92.04

Figure 7 Dataset Size Vs. Recall



	10K	50K	100K	150K	200K
IoBTSec-RPL	0.98	0.97	0.972	0.955	0.96
LGBM	0.92	0.933	0.931	0.942	0.935
LSTM	0.869	0.868	0.736	0.738	0.871
DBN	0.862	0.873	0.868	0.869	0.865

Figure 8 Dataset Size Vs MCC

6. CONCLUSION

This paper proposed a hybrid DL-based multi-attack detection and classification algorithm named IoBTSec-RPL for RPL-IoBT. The proposed IoBTSec model detected and classified the rank, version, blackhole, hello flood, and sinkhole attacks. Initially, the EPOA removes irrelevant features from network traffic to minimize computational complexity and improve classification accuracy. The ACGAN-based class-specific augmented network traffic facilitated the multiple attack classification in the imbalanced class environment, even detecting novel attack patterns without over fitting. Further, the hybrid of LSTM with DBN improved multiclass attack detection accuracy and the military surveillance performance in IoBT. Finally, the experimental results proved that the proposed IoBTSec-RPL outperformed the comparative baseline models and yielded 98.3% recall. The proposed IoBTSec-RPL improves accuracy by 2.16%, 5.73%, and 6.06% than the LGBM, LSTM, and DBN, respectively, for 200K traffic samples. Privacy-preserving distributed machine learning, such as federated learning, is highly recommended for future work in an IoBT-based edge network environment for attack detection models. Network vulnerability analysis using fuzzing methods in the IoBT environment is also recommended as a future research direction.

REFERENCES

[1] Nord, J. H., Koohang, A., & Paliszkiewicz, J. (2019). The Internet of Things: Review and theoretical framework. Expert Systems with Applications, 133, 97–108. https://doi.org/10.1016/j.eswa.2019.05.014

[2] Farooq, M. J., & Zhu, Q. (2018). On the secure and reconfigurable multi-layer network design for critical information dissemination in the internet of battlefield things (IoBT). IEEE Transactions on Wireless Communications, 17(4), 2618–2632. https://doi.org/10.1109/twc.2018.2799860

[3] Sobral, J. V., Rodrigues, J. J., Rabêlo, R. A., Al-Muhtadi, J., &Korotaev, V. (2019). Routing protocols for low power and lossy networks in Internet of things applications. Sensors, 19(9), 2144. https://doi.org/10.3390/s19092144.

[4] Ekpenyong, M. E., Asuquo, D. E., Udo, I. J., Robinson, S. A., & Ijebu, F. F. (2022). IPv6 routing protocol enhancements over low-power and lossy networks for IoT applications: A systematic review. New Review of Information Networking, 27(1), 30–68. https://doi.org/10.1080/13614576.2022.2078396

[5] Kharrufa, H., Al-Kashoash, H. A. A., & Kemp, A. H. (2019). RPL-based routing protocols in IoT applications: A review. IEEE Sensors Journal, 19(15), 5952–5967. https://doi.org/10.1109/jsen.2019.2910881

[6] Simoglou, G., Violettas, G., Petridou, S., & Mamatas, L. (2021). Intrusion detection systems for RPL security: A comparative analysis. Computers & Security, 104(102219), 102219. https://doi.org/10.1016/j.cose.2021.102219.

[7] Raoof, A., Matrawy, A., & Lung, C.-H. (2019). Routing attacks and mitigation methods for RPL-based internet of things. IEEE Communications Surveys & Tutorials, 21(2), 1582–1606. https://doi.org/10.1109/comst.2018.2885894

[8] Pasikhani, A. M., Clark, J. A., Gope, P., & Alshahrani, A. (2021). Intrusion detection systems in RPL-based 6LoWPAN: A systematic literature review. IEEE Sensors Journal, 21(11), 12940–12968. https://doi.org/10.1109/jsen.2021.3068240

[9] Al-Amiedy, T. A., Anbar, M., Belaton, B., Kabla, A. H. H., Hasbullah, I. H., & Alashhab, Z. R. (2022). A systematic literature review on machine and Deep Learning approaches for detecting attacks in RPL-based 6LoWPAN of internet of things. Sensors (Basel, Switzerland), 22(9), 3400. https://doi.org/10.3390/s22093400

[10] Thakkar, A., & Lohiya, R. (2021). A review on machine learning and deep learning perspectives of IDS for IoT: Recent updates, security issues, and challenges. Archives of Computational Methods in

Engineering. State of the Art Reviews, 28(4), 3211–3243. https://doi.org/10.1007/s11831-020-09496-0

[11] Zahra, F., Jhanjhi, N.Z., Brohi, S.N., Khan, N.A., Masud, M., &Alzain, M.A. (2022). Rank and Wormhole Attack Detection Model for RPL-Based Internet of Things Using Machine Learning. Sensors (Basel, Switzerland), 22(18), 6765. https://doi.org/10.3390/s22186765

[12] Zahra, F., Jhanjhi, N. Z., Khan, N. A., Brohi, S. N., Masud, M., & Aljahdali, S. (2022). Protocol-specific and sensor network-inherited attack detection in IoT using machine learning. Applied Sciences (Basel, Switzerland), 12(22), 11598. https://doi.org/10.3390/app122211598

[13] Osman, M., He, J., Mokbal, F. M. M., Zhu, N., & Qureshi, S. (2021). ML-LGBM: A machine learning model based on light gradient boosting machine for the detection of version number attacks in RPL-based networks. IEEE Access: Practical Innovations, Open Solutions, 9, 83654–83665. https://doi.org/10.1109/access.2021.3087175

[14] Raghavendra, T., Anand, M., Selvi, M., Thangaramya, K., Santhosh Kumar, S. V. N., & Kannan, A. (2022). An Intelligent RPL attack detection using Machine Learning-Based Intrusion Detection System for Internet of Things. Procedia Computer Science, 215, 61–70. https://doi.org/10.1016/j.procs.2022.12.007

[15] Verma, A., & Ranga, V. (2019). ELNIDS: Ensemble Learning based Network Intrusion Detection System for RPL based Internet of Things. 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU). " In 2019 4th International conference on Internet of Things: Smart innovation and usages (IoT-SIU), pp. 1-6. https://doi.org/10.1109/IoT-SIU.2019.8777504.

[16] Yaakoubi, Faicel, AymenYahyaoui, WadiiBoulila, and RabahAttia. "An XGBoost-Based Approach for an Efficient RPL Routing Attack Detection." In Computational Collective Intelligence: 14th International Conference, ICCCI 2022, Hammamet, Tunisia, September 28–30, 2022, Proceedings, pp. 611-623. https://doi.org/10.1007/978-3-031-16014-1_48.

[17] Bouazza, A., Debbi, H., &Lakhlef, H. (2022). Machine Learning-based Intrusion Detection System Against Routing Attacks in the Internet of Things. Proceedings http://ceur-ws. org ISSN, 1613, 0073.

[18] Cakir, S., Toklu, S., & Yalcin, N. (2020). RPL attack detection and prevention in the internet of things networks using a GRU based deep learning. IEEE Access: Practical Innovations, Open Solutions, 8, 183678–183689. https://doi.org/10.1109/access.2020.3029191.

[19] Morales-Molina, C. D., Hernandez-Suarez, A., Sanchez-Perez, G., Toscano-Medina, L. K., Perez-Meana, H., Olivares-Mercado, J., Portillo-Portillo, J., Sanchez, V., & Garcia-Villalba, L. J. (2021). A Dense Neural Network approach for detecting Clone ID attacks on the RPL protocol of the IoT. Sensors (Basel, Switzerland), 21(9), 3173. https://doi.org/10.3390/s21093173.

[20] Choukri, W., Lamaazi, H., & Benamar, N. (2020). RPL rank attack detection using Deep Learning. 2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT), 1-6. https://doi.org/10.1109/3ICT51146.2020.9311983.

[21] Albishari, M., Li, M., Zhang, R., & Almosharea, E. (2023). Deep learning-based early stage detection (DL-ESD) for routing attacks in Internet of Things networks. The Journal of Supercomputing, 79(3), 2626–2653. https://doi.org/10.1007/s11227-022-04753-4

[22] Saba, T., Rehman, A., Sadad, T., Kolivand, H., & Bahaj, S. A. (2022). Anomaly-based intrusion detection system for IoT networks through deep learning model. Computers & Electrical Engineering: An International Journal, 99(107810), 107810. https://doi.org/10.1016/j.compeleceng.2022.107810.

[23] Shahnawaz Ahmad, M., & Mehraj Shah, S. (2022). Unsupervised ensemble based deep learning approach for attack detection in IoT network. Concurrency and Computation: Practice & Experience, 34(27). https://doi.org/10.1002/cpe.7338.

[24] Nayak, S., Ahmed, N., & Misra, S. (2021). Deep learning-based reliable routing attack detection mechanism for industrial internet of things. Ad Hoc Networks, 123(102661), 102661. https://doi.org/10.1016/j.adhoc.2021.102661

[25] Shao, S., Wang, P., & Yan, R. (2019). Generative adversarial networks for data augmentation in machine fault diagnosis. Computers in Industry, 106, 85–93. https://doi.org/10.1016/j.compind.2019.01.001.

[26] Trojovský, P., & Dehghani, M. (2022). Pelican Optimization Algorithm: A novel nature-inspired algorithm for engineering applications. Sensors (Basel, Switzerland), 22(3), 855. https://doi.org/10.3390/s22030855.

Authors

Mrs. K. Kowsalyadevi is currently pursuing her Ph.D in Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India. Her area of interests includes Wireless Sensor Networks, Artificial Intelligence and the Internet of Things.

Dr. N.V. Balaji, presently working as a Dean, Arts, Science, Commerce and Management, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India. As a veteran educationalist for more than a decade he renders his service as an academic contriver with zeal for 21 years. He enriched training and placements with immense enthusiasm scaling the academy to great heights. Dr. Balaji a pertinacious personality brought laurels upon the institutions by representing it at Cambridge University for Business English Certifications. He is an honourable recipient of the award of Ambassador for Computer based Learning and Assessment category in the year 2015. Adding to his reputation he has accumulated 15 years of research experience in the field of computer science. He exhibits deep interest towards various genres such as Neural Networks, Fuzzy Logic, Image Processing, Classification and Data Mining.

How to cite this article:

K. Kowsalyadevi, N.V.Balaji, “IoBTSec-RPL: A Novel RPL Attack Detecting Mechanism Using Hybrid Deep Learning Over Battlefield IoT Environment”, International Journal of Computer Networks and Applications (IJCNA), 10(4), PP: 637-650, 2023, DOI: 10.22247/ijcna/2023/223317.