The Scientific Temper (2025) Vol. 16 (3): 3856-3862
E-ISSN: 2231-6396, ISSN: 0976-8653
Doi: 10.58414/SCIENTIFICTEMPER.2025.16.3.02
https://scientifictemper.com/

RESEARCH ARTICLE

RFSVMDD: Ensemble of multi-dimension random forest and custom-made support vector machine for detecting RPL DDoS attacks in an IoT-based WSN environment

R. Sakthiraman, L. Arockiam

Abstract

Growing dependence on the internet of things (IoT) and wireless sensor networks (WSNs) has led to critical security issues, especially concerning distributed denial of service (DDoS) attacks based on RPL. Such attacks can severely compromise the network's security, reliability, and efficiency. To effectively address this problem, this research proposes (RFSVMDD) a novel hybrid detection model that combines a multi-dimensional random forest (MDRF) with a custom-made support vector machine (CSVM). The proposed technique uses MDRF to provide scalability for consistent feature selection and anomaly detection across high-dimensional datasets. CSVM reduces false positives and increases detection accuracy through its improved classification of potential threats. Experimental assessments in simulated IoT-based WSN environments show that the model outperforms conventional machine learning methods regarding accuracy, detection speed, and durability. This novel ensemble approach presents a promising solution by enhancing IoT and WSN networks against RPL DDoS attacks.

Keywords: Internet of Things, Wireless sensor networks, Security, Distributed denial of service attacks, Routing protocol for low power and lossy networks.

Introduction

In recent years, the proliferation of Internet of Things (IoT) devices has revolutionized various domains, ranging from smart homes to industrial automation. However, the inherent vulnerabilities of IoT networks, particularly those employing wireless sensor networks (WSNs) and the routing protocol for low-power and lossy networks (RPL), have rendered them susceptible to security threats (Azzedin *et al*., 2023).

Department of Computer Science, St. Joseph's College (Autonomous), (Affiliated to Bharathidasan University), Tiruchirappalli – 620002, India.

***Corresponding Author:** R. Sakthiraman, Department of Computer Science, St. Joseph's College (Autonomous), (Affiliated to Bharathidasan University), Tiruchirappalli – 620002, India, E-Mail: chandransakthi887@gmail.com

How to cite this article: Sakthiraman, R., Arockiam, L. (2025). RFSVMDD: Ensemble of multi-dimension random forest and custom-made support vector machine for detecting RPL DDoS attacks in an IoT-based WSN environment. The Scientific Temper, **16**(3):3856-3862.

Doi: 10.58414/SCIENTIFICTEMPER.2025.16.3.02

Source of support: Nil

Conflict of interest: None.

The IPv6 RPL is designed to organize resource-constrained nodes in a destination-oriented directed acyclic graph (DODAG) topology. It supports multi-hop communication and is optimized for many-to-one and one-to-one message exchanges. Among these threats, distributed denial of service (DDoS) attacks targeting RPL-based IoT networks pose a significant challenge due to their potential to disrupt critical services and compromise network integrity (Ahmim *et al*., 2023). Smart device integration increases security vulnerabilities while automating regular operations. According to Microsoft, DDoS attacks are a serious menace that has increased by 300%.

To mitigate the risks associated with RPL DDoS attacks in IoT-based WSN environments, this paper proposes an innovative ensemble approach combining a multi-dimensional random forest (MDRF) and a custom-made support vector machine (SVM).

Existing Methods

(Zhou *et al*., 2022) This article addresses the WD-DNS model is a multidimensional fusion model that uses feature extraction from resolutions, requests, and domain names to detect deep DNS attacks. By using deep learning, the model outperforms single-dimensional detection algorithms in terms of accuracy. One drawback is its dependence on

Received: 09/02/2025 **Accepted:** 21/02/2025 **Published** : 20/03/2025

pre-set feature extraction, which might not be long-term enough to capture dynamic patterns of new threats.

(Alashhab *et al*., 2024) In this work, an ensemble online machine learning (OML) model utilizing BernoulliNB, passive-aggressive, SGD, and MLP classifiers is proposed for the real-time detection of DDoS attacks in software-defined networks (SDN). The model continuously updates with new traffic data, providing improved flexibility against dynamic threats such as zero-day and low-rate attacks. The intrusion prevention system (IPS) and intrusion detection system (IDS) are integrated into the model to improve detection and mitigation. Results from experiments indicate improved mitigation performance and precision. Inefficient feature selection and controller overhead are still problems.

(Wang *et al*., 2022) This paper suggests an ensemble online machine learning (OML) model that uses BernoulliNB, passive-aggressive, SGD, and MLP classifiers to detect DDoS attacks in real time for software-defined networks (SDN). As new traffic data is continuously added, the model provides increased flexibility against dynamic threats such as zero-day and low-rate attacks. The paradigm combines intrusion prevention and detection systems (IDS and IPS) to improve detection and mitigation. Results from experiments show improved accuracy and mitigation capabilities. Open problems include things like controller overhead and effective feature selection.

(Azzedin *et al*., 2023) The author addresses vulnerabilities of RPL-based IoT networks that are examined in this paper from the standpoint of version number tampering and hello flooding. In order to minimize malicious activity, it offers a trust-based mitigation method that assesses node behavior. Simulations using Contiki and Cooja are used to show how well the solution performs in terms of lowering energy drain and network recovery. Whitewashing and trust erosion are still issues that can jeopardize dependability in the long run.

(Ahmed *et al*., 2023) This research advocates the implementation of an MLP classification model on internal datasets to pinpoint application layer DDoS attacks. Drawing upon features extracted from incoming network traffic, characterized by significant variances, the research explores various groupings of attack attributes and devises a framework to discern between assailants, suspects, and legitimate users.

(Ullah *et al*., 2023) This article presents a machine learning-based technique for identifying DDoS attacks, consisting of three modules: preprocessing, attribute selection, and a detection and prevention system. Technically, the incoming traffic attributes are first normalized on a standard scale during the preprocessing stage. The random forest approach was then used to identify the 30 most productive characteristics from 79. The technique has been evaluated on a widely accessible dataset, and the result was evaluated by the accuracy.

(Bhale *et al*., 2023) This study introduces OPTIMIST, a distributed intrusion detection system for Internet of Things networks that can identify both high- and low-rate DDoS attacks. For better detection, it integrates with an LSTM model trained using flows generated using Wasserstein GAN and uses a weighted minimum vertex cover technique to locate the IDS node appropriately. The capacity to balance energy usage and detection performance is confirmed by experimental results. The current IDS solutions are energy inefficient, non-transparent, and either high-rate or low-rate oriented.

(Albishari *et al*., 2023) This article introduces DL-ESD, a deep learning-based intrusion detection model designed to protect IoT networks from RPL routing attacks like version number alteration, hello flood, and lowered rank. The model combines multi-layer perceptron (MLP) and deep neural networks (DNN) to achieve a classification accuracy of over 98.98% based on the IRAD dataset. For improved detection performance, it combines feature selection, normalization, and dimensionality reduction (LDA). Inefficient parallel processing capabilities and control message overhead in resource-constrained IoT devices are obstacles to current solutions. In the future, edge computing will be used to improve detection scalability and provide real-time defense against assaults.

(Alsukayti *et al*., 2022) This paper, proposes CDRPL, a distributed and lightweight detection and mitigation method based on RPL for Version Number (VN) threats against Internet-of-things (IoT) networks. Better security with less overhead and external entities is provided by CDRPL, which requires fewer extensions than RPL. In comparison to existing solutions, it provides robust network performance, rapid topology convergence, and fast attack detection. The complexity, high processing demands, and delayed convergence of current VN attack mitigation techniques make them unsuitable for IoT devices with limited resources.

3. Proposed Method

The proposed work titled "RFSVMDD: Ensemble of multi-dimension random forest and custom-made support Vector Machine for detection RPL DDoS attacks in IoT based WSN environment" is the integration of two functional modules namely multi-dimension random forest, and custom-made support vector machine. Multi-dimension random forest is used to identify suspicious RPL-based DDoS attacks in a swifter manner. The custom-made support vector machine is used to make the final classification of DDoS attacks with the bias provided by the prior module. A comprehensive description of these modules is provided in this section.

3.1. Multi-Dimension Random Forest (MDRF)

The number of features in a typical network transaction depends on several key factors such as communication

protocol, routing protocol, network architecture, and the hardware components involved in the communication. Let n_f be the number of features in a captured network transaction, $\sqrt{n_f}$ is the number of sample features taken in a regular random forest model during the bootstrapping phase. The features are maintained in a set $F = \left\{f_1, f_2 \cdots f_{n_f}\right\}$. Similarly, the feature groups are maintained in a set $G = \left\{g_1, g_2 \ldots g_{n_g}\right\}$. In contrast to the standard random forest model, MDRF takes all the input features and splits them into n_g number of separate groups as in equation 1 to improve the true positives, true negatives and accuracy.

$$n_g = \frac{n_f}{2 \times \sqrt{n_f}} \qquad \text{Equation (1)}$$

Therefore, the number of features will be $2 \times \sqrt{n_f}$ for every group. Padding is applied from the beginning of the feature set to the last group if the number of members of the group is less than $2 \times \sqrt{n_f}$ as illustrated in Figure 1.

Every group $g_x \in G$ is considered as a random forest in different dimensions, thus there are n_g dimensions maintained in the MDRF model as in Equation 2.

$$R = \left\{r_1, r_2, \ldots r_{n_g}\right\} \qquad \text{Equation (2)}$$

Each dimension random forest r_x is bootstrapped with their own set of decision trees as in Equation 3.

$$r_x = \left\{t_1, t_2, \ldots t_n\right\} \qquad \text{Equation (3)}$$

The dimension reduction process is performed as per the procedure given in MDRF dimension reduction algorithm.

Algorithm 1: MDRF-DR

- *Input*

Multi-dimension random forest Set R, Wireshark-assisted hypervisor input

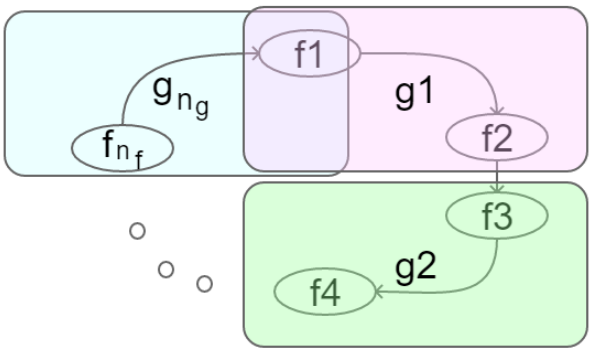

Figure 1: Features and Group padding

- *Output*

Dimension reduced random forest set R^*

Step 1: Load R

Step 2: Let η be the number of IoT-WSN nodes in the network

Step 3: Let $B = \left\{\beta_{r_1}, \beta_{r_2} \cdots \beta_{r_{n_g}}\right\}$ be the set of beneficial score of individual random forests

Step 4: Initialize beneficial scores as $\forall \mathrm{i} = 1 \rightarrow n_g := \beta_{\gamma_i} = 0$

Step 5: Fetch Hypervisor DDoS flag for every IoT-WSN node

Step 6: For $\mathrm{i} = 0 \rightarrow \eta$

Step7: $\forall j = 1 \rightarrow n_9 := \beta_{r_i} + \begin{cases} 1 if\ Result(r_i) \equiv Result(Hypervisor) \\ 0 otherwise \end{cases}$

Step 8: End // For i

Step 9: Initialize Dimension count $\Delta_c = 0$

Step 10: For $\mathrm{i} = 1 \rightarrow n_g$

Step 11: Find Selection Status δ of r_i by Equation 4

Step 12: If $\delta = TRUE$ then

Step 13: Add $r_i \in R \rightarrow R^*$ as r_{Δ_c}

Step 14: Increment Δ_c by 1

Step 15: End If // δ

Step 16: End For // i

Step 17: Return R^*

The Selection Status Equation is as follows

$$\delta = \begin{cases} TRUE\ if\ \beta_i > \dfrac{\eta}{2} \\ FALSE\ otherwise \end{cases} \qquad \text{Equation (4)}$$

By this way, the MDRF-DR algorithm reduces the input multi-dimension random forests by eliminating the members which are below 50% beneficial. The beneficial Multi-Dimension Random Forest set R^* is propelled to the successive functional module.

Custom-made Support Vector Machine (CSVM)

Supervised learning is used in the CSVM model. Though there is no labeled dataset used, the Wireshark-assisted hypervisor output provides the DDoS attack labels for every network transaction that occurs in the specific IoT-WSN network environment. The feature selection process of CSVM is inherited by the MDRF model. It is spontaneous that the beneficial random forests have a higher probability of holding the superior impact features among all the input features. Let $\sigma = \varnothing$ be the feature set to be used by the CSVM module. The selected features are included to σ by equation 5.

$$\forall \mathrm{i} = 1 \rightarrow \Delta_c := \sigma = \sigma \cup Features(g_i \in G) \qquad \text{Equation (5)}$$

Let n_σ be the number of features selected to be processed with CSVM, then a $n_\sigma - 1$ dimensional hyperplane is initialized for CSVM. All possible combinations of X-Y axis features are leveraged by the CSVM Hyperplane Selection algorithm to improve the true positive rate (TPR) and false positive rate (FPR).

Algorithm 2: CSVM-HS

- *Input*

σ

- *Output*

Possible hyperplane vector features

Step 1: Load σ

Step 2: Initialize hyperplane set $H = \varnothing$

Step 3: Initialize hyperplane count $c = 0$

Step 4: For $\mathrm{i} = 1 \rightarrow n_\sigma$

Step 5: For $j = i + 1 \rightarrow n_\sigma - 1$

Step 6: If $\mathrm{i} \neq j$, then

Step 7: Increment c by 1

Step 8: Add $h_c = \left(f_i, f_j\right)$

Step 9: End if // $\mathrm{i} \neq j$

Step 10: End if // j

Step 11: End if // i

Step 12: Return H

All hyperplanes are trained using non-linear kernels. The overall DDoS detection process is accomplished by the RFSVMDD Algorithm.

Algorithm 3: RFSVMDD

- *Input*

Input network transaction

- *Output*

DDoS flag

Step 1: Let ω_R be the output of MDRF

Step 2: Let ω_S be the output of CSVM

Step 3: Let ω_Q be the normalized Output Quotient

Step 4: Read network transaction

Step 5: $\omega_R = \frac{1}{\Delta_c} \sum_{i=1}^{\Delta_c} Result\left(r_i\right)$

Step 6: $\omega_S = \frac{1}{n_\sigma - 1} \sum_{i=1}^{n_\sigma - 1} Result\left(h_i\right)$

Step 7: Compute $\omega_Q = \omega_R + \left(\frac{\omega_S - \omega_R}{2}\right)$

Step 8: $Output = \begin{cases} DDoS \; if \; \omega_Q > \frac{1}{2} \\ Not \; DDos \; otherwise \end{cases}$

Step 9: Go to Step 4

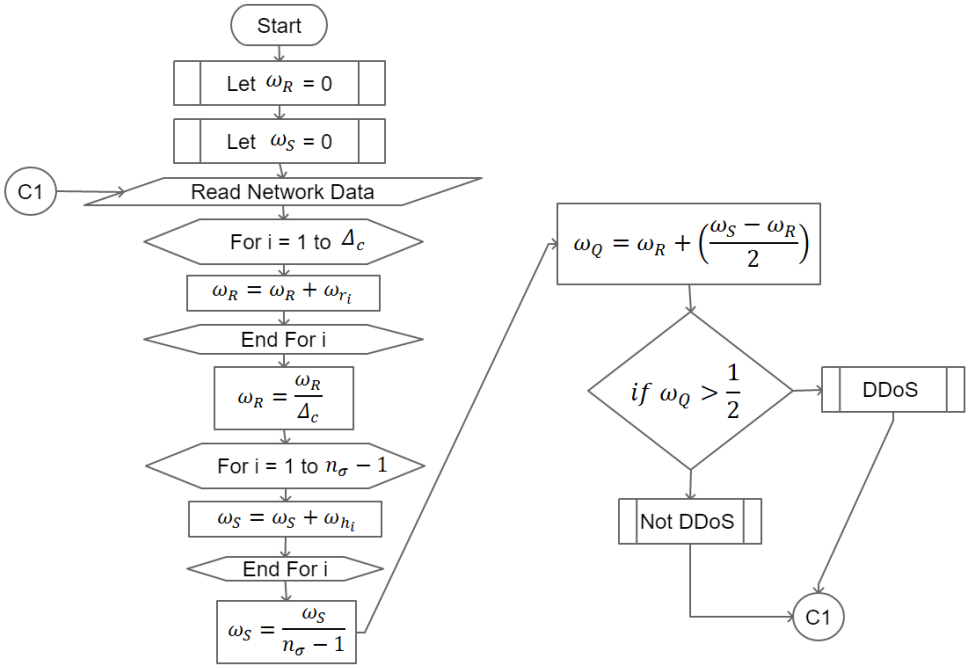

Figure 2: RFSVMDD Flowchart

A flowchart of RFSVMDD is given in Figure 2

Through this process, the integration of MDRF and CSVM methods is capable of detecting RPL DDoS attacks in IoT-WSN environments.

Experimental Setup

A computer with an i7 processor, 16GB of Memory, and 1TB of storage is used to implement the proposed method. A server is leashed from i2k2.com to get the real-time IoT-WSN data and to perform the DDoS detection process. The server is equipped with an industrial standard Wireshark network tracing tool to evaluate the performance of the different methods discussed here. A dedicated user interface (UI) is designed to load the programming scripts to the server, and the UI is also capable of interacting with Wireshark and the server to fetch the performance metrics. Visual Studio IDE is used to develop the UI. C++ 20.0 programming language is used to develop the network scripts and snippets to communicate with the Wireshark network monitoring software. Since the entire implementation uses real-time network data, the implication of any pre-created dataset is extraneous. The performance during the training phase is logged for every 7% data till 70%, and during the testing phase, the performance metrics are logged for every 3% data till 30%. Therefore, a 70:30 training-testing ratio is followed during the experiments.

Results and Analysis

The Tables 1 to 6 show the outcomes of the proposed RFSVMDD method approach in comparison to other existing models, such as deep learning early-stage detection (DLESD), deep learning algorithm multilayer perceptron (DLMLP), machine Learning based dynamic attributes selection technique (MLDAST), lightweight and transparent intrusion detection system (LTIDS). Demonstrates a noteworthy improvement in all evolution metrics, including F-score, G-mean, specificity, accuracy, precision, and sensitivity.

When compared to existing methods, the proposed RFSVMDD method has obtained 98.91% accuracy, as Table 1 illustrates. Table 2 shows that the RFSVMDD approach outperformed the precision of 98.81%. In the Sensitivity, that RFSVMDD receives a 99.00%. Table 4 compares specificity achieved at 98.82%, while Table 5 F score metrics reach 98.91%. Ultimately, the G-mean value of 98.91% is shown in Table 6.

Table 1: Comparison of accuracy (%)

Data (%)	*DLESD*	*DLMLP*	*MLDAST*	*LTIDS*	*RFSVMDD*
3	96.47	98.08	97.27	97.27	98.92
6	96.39	98.09	97.30	97.17	98.96
9	96.46	98.05	97.28	97.17	98.88
12	96.46	98.07	97.31	97.20	98.94
15	96.41	98.08	97.27	97.25	98.92
18	96.47	98.13	97.35	97.25	98.90
21	96.48	98.07	97.24	97.24	98.95
24	96.50	98.01	97.32	97.24	98.96
27	96.39	98.09	97.32	97.19	98.94
30	96.39	98.07	97.26	97.20	98.91

Table 2: Comparison precision (%)

Data (%)	*DLESD*	*DLMLP*	*MLDAST*	*LTIDS*	*RFSVMDD*
3	97.03	98.18	96.00	96.51	98.84
6	96.92	98.24	95.89	96.37	98.89
9	96.95	98.13	96.00	96.40	98.79
12	96.90	98.16	95.94	96.35	98.79
15	96.98	98.27	96.00	96.45	98.79
18	96.92	98.27	96.00	96.40	98.73
21	96.95	98.18	95.94	96.37	98.89
24	97.06	98.10	96.05	96.40	98.87
27	96.92	98.27	96.00	96.40	98.87
30	96.90	98.27	95.97	96.40	98.81

Table 3: Comparison of sensitivity (%)

Data (%)	*DLESD*	*DLMLP*	*MLDAST*	*LTIDS*	*RFSVMDD*
3	95.95	97.98	98.50	98.00	99.00
6	95.90	97.95	98.66	97.94	99.03
9	96.00	97.98	98.53	97.92	98.98
12	96.05	97.98	98.64	98.02	99.08
15	95.90	97.90	98.50	98.02	99.06
18	96.05	98.01	98.67	98.08	99.06
21	96.05	97.95	98.50	98.08	99.01
24	95.98	97.92	98.56	98.05	99.06
27	95.90	97.93	98.61	97.94	99.01
30	95.92	97.88	98.50	97.97	99.00

Table 4: Comparison of specificity (%)

Data (%)	*DLESD*	*DLMLP*	*MLDAST*	*LTIDS*	*RFSVMDD*
3	97.00	98.18	96.10	96.56	98.84
6	96.89	98.23	96.00	96.43	98.90
9	96.92	98.13	96.10	96.45	98.79
12	96.87	98.15	96.05	96.41	98.79
15	96.94	98.26	96.10	96.51	98.79
18	96.90	98.26	96.10	96.46	98.74
21	96.92	98.18	96.05	96.43	98.90
24	97.02	98.10	96.15	96.46	98.87
27	96.89	98.26	96.10	96.46	98.87
30	96.86	98.26	96.07	96.46	98.82

Table 5: Comparison of F-score (%)

Data (%)	*DLESD*	*DLMLP*	*MLDAST*	*LTIDS*	*RFSVMDD*
3	96.49	98.08	97.23	97.25	98.92
6	96.41	98.10	97.26	97.15	98.96
9	96.47	98.05	97.25	97.15	98.88
12	96.47	98.07	97.27	97.18	98.94
15	96.43	98.08	97.23	97.23	98.92
18	96.48	98.14	97.31	97.23	98.89
21	96.50	98.07	97.21	97.22	98.95
24	96.52	98.01	97.29	97.22	98.96
27	96.41	98.10	97.29	97.16	98.94
30	96.41	98.07	97.22	97.18	98.91

Table 6: Comparison of G-mean (%)

Data (%)	*DLESD*	*DLMLP*	*MLDAST*	*LTIDS*	*RFSVMDD*
3	96.47	98.08	97.29	97.28	98.92
6	96.39	98.09	97.32	97.18	98.96
9	96.46	98.05	97.31	97.18	98.88
12	96.46	98.07	97.34	97.21	98.94
15	96.42	98.08	97.29	97.26	98.92
18	96.47	98.13	97.38	97.26	98.90
21	96.49	98.07	97.27	97.25	98.95
24	96.50	98.01	97.35	97.25	98.96
27	96.39	98.09	97.35	97.20	98.94
30	96.39	98.07	97.28	97.21	98.91

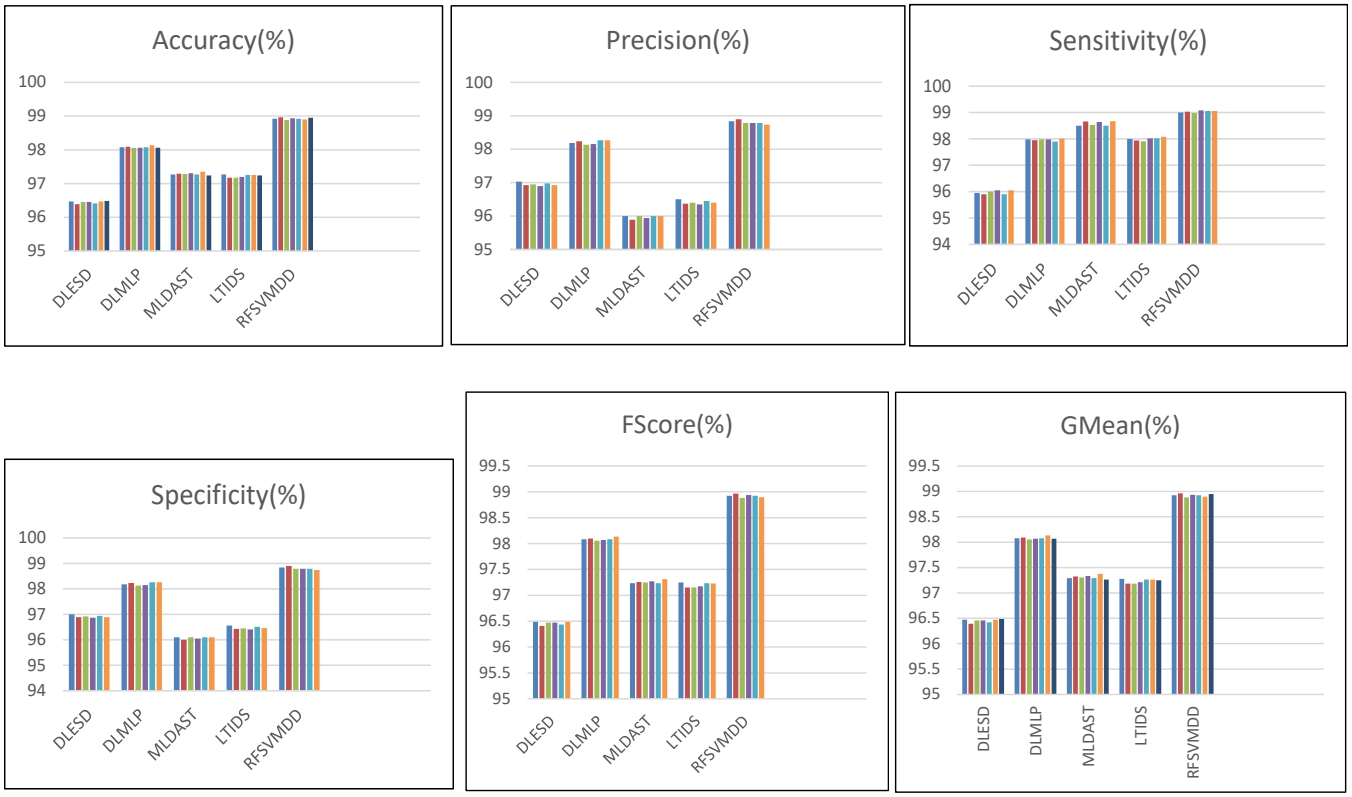

Figure 3: Overall comparative results of RFSVMDD

Figure 3 shows the graphical representation outcomes of the proposed RFSVMDD method approach in comparison to other existing models, such as deep learning early stage detection (DLESD), deep learning algorithm multilayer perceptron (DLMLP), machine learning-based dynamic attributes selection technique (MLDAST), lightweight and transparent intrusion detection system (LTIDS). The graph above makes it obvious that the RFSVMDD approach outperforms the current approaches in every assessment metric.

Conclusion

This research presents a sophisticated and effective solution for detecting RPL DDoS attacks in IoT-based WSN environments. This work has achieved notable detection accuracy and robustness advancements by harnessing the collective power of an ensemble of multi-dimensional random forests and a custom-made support vector machine (RFSVMDD). The experimental results validate the efficacy of our approach, showcasing its ability to accurately discern RPL DDoS attacks within complex network traffic.

This research contributes to the advancement of security measures in IoT systems, offering a reliable defense mechanism against emerging threats. Future work may involve further optimization of the proposed algorithms and their deployment in real-world IoT environments to strengthen cybersecurity in the ever-expanding IoT landscape.

References

Ahmed, S., Khan, Z. A., Mohsin, S. M., Latif, S., Aslam, S., Mujlid, H., ... & Najam, Z. (2023). Effective and efficient DDoS attack detection using deep learning algorithm, multi-layer perceptron. *Future Internet*, *15*(2), 76.

Ahmim, A., Maazouzi, F., Ahmim, M., Namane, S., & Dhaou, I. B. (2023). Distributed denial of service attack detection for the Internet of Things using hybrid deep learning model. *IEEE Access*, *11*, 119862-119875.

Alashhab, A. A., Zahid, M. S., Isyaku, B., Elnour, A. A., Nagmeldin, W., Abdelmaboud, A., ... & Maiwada, U. (2024). Enhancing DDoS attack detection and mitigation in SDN using an ensemble online machine learning model. *IEEE Access*.

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.

Alsukayti, I. S., & Singh, A. (2022). A lightweight scheme for mitigating RPL version number attacks in IoT networks. *IEEE Access*, *10*, 111115-111133.

Azzedin, F. (2023). Mitigating denial of service attacks in RPL-based IoT environments: trust-based approach. *IEEE Access*, *11*, 129077-129089.

Bhale, P., Chowdhury, D. R., Biswas, S., & Nandi, S. (2023). OPTIMIST: lightweight and transparent IDS with optimum placement strategy to mitigate mixed-rate DDoS attacks in IoT networks. *IEEE Internet of Things Journal*, *10*(10), 8357-8370.

Ullah, S., Mahmood, Z., Ali, N., Ahmad, T., & Buriro, A. (2023). Machine learning-based dynamic attribute selection technique for DDoS attack classification in IoT networks. *Computers*, *12*(6), 115.

Wang, J., & Wang, L. (2022). SDN-defend: a lightweight online attack detection and mitigation system for DDoS attacks in SDN. *Sensors*, *22*(21), 8287.

Zhou, Y., Yang, L., Wang, Z., Li, G., & Ning, X. (2022, December). DNS attack detection based on multi-dimensional fusion model. In *2022 International Conference on Networking and Network Applications (NaNA)* (pp. 74-81). IEEE.