

Adaptive intrusion detection system with DBSCAN to enhance banking cybersecurity

Sathiyaseelan Periyasamy¹, Anubhav Kumar², Karupusamy Muthulakshmi³,
Thenmozhi Elumalai³, Prabu Kaliyaperumal², Rajakumar Perumal⁴

¹Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai, India

²School of Computer Science and Engineering, Galgotias University, Greater Noida, India

³Department of Information Technology, Panimalar Engineering College, Chennai, India

⁴Department of Computer Science and Applications, Sharda School of Computing Science and Engineering, Sharda University, Greater Noida, India

Article Info

Article history:

Received Jan 14, 2025

Revised Jul 2, 2025

Accepted Aug 6, 2025

Keywords:

Anomaly detection

Autoencoder

CICIDS2017

CSECICIDS2018

Cyber threats

Multi-class classification

Securing bank

ABSTRACT

The accelerating pace of digital transformation in the banking sector has highlighted the critical need for comprehensive cybersecurity strategies capable of countering evolving cyber threats. This study introduces an innovative intrusion detection framework tailored for banking environments, leveraging the CICIDS2017 and CSECICIDS2018 datasets for evaluation and validation. The proposed framework integrates data preprocessing, feature reduction, and advanced attack detection methods to enhance detection accuracy. A basic autoencoder is utilized for dimensionality reduction, streamlining input data while preserving essential attributes. The density-based spatial clustering of applications with noise (DBSCAN) algorithm is then applied for attack detection, enabling the detection of intricate attack patterns and their classification into specific attack groups. The proposed adaptive intrusion detection system (IDS) framework demonstrates outstanding performance, achieving precision, recall, F1-score, and accuracy rates exceeding 98%. Comparative evaluations against conventional techniques, such as support vector machines (SVM), long short-term memory (LSTM), and K-means, highlight its superiority in terms of accuracy and computational efficiency. This research addresses key challenges, including high-dimensional datasets, class imbalance, and dynamic threat landscapes, offering a scalable and efficient solution to enhance the security of banking operations and enable proactive threat mitigation in the sector.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Prabu Kaliyaperumal

School of Computer Science and Engineering, Galgotias University

Greater Noida, Delhi NCR, India

Email: mega.prabu@gmail.com

1. INTRODUCTION

The growing reliance on digital infrastructure in the banking industry has heightened its vulnerability to cyber threats [1], [2]. Cybercriminals exploit vulnerabilities to compromise sensitive financial data, disrupt services and erode the confidence of customers [3]. As financial institutions increasingly adopting interconnected systems, addressing cybersecurity challenges has become a pressing priority [4]. This research introduces a holistic intrusion detection framework aimed at strengthening cybersecurity measures in banking operations. The banking sector bears a critical obligation to safeguard customer assets while complying with rigorous regulatory requirements. Cyber threats against banking

systems have become increasingly advanced, encompassing tactics like malware, phishing schemes, insider threats, and advanced persistent threats (APTs) [5], [6]. Such attacks jeopardize confidential information and interrupt operations, leading to substantial financial losses and a decline in organizational credibility. The emergence of multi-vector threats underscores the critical need for sophisticated security solutions that can identify, assess, and neutralize both existing and novel threats in real-time. Traditional intrusion detection systems (IDS) typically utilize signature-based methods, which perform well against established threats but face challenges in identifying new or evolving attack patterns [7], [8]. Moreover, high-dimensional and imbalanced datasets intensify the complexities of threat detection, leading to an increased occurrence of false positives and undetected attacks [9], [10]. This necessitates the development of advanced, data-driven approaches that leverage machine learning and deep learning technologies to enhance the security of banking systems.

In recent years, numerous approaches have been explored to address these challenges. For instance, Gheni and Yaseen [11] introduced a two-step clustering-based intrusion detection model utilizing the CICIoT2023 dataset. Leveraging gaining-sharing knowledge (GSK) optimization and multilayer perceptron (MLP), it achieved 99.26% accuracy and a 62.45% dataset size reduction, enhancing efficacy and speed. However, reliance on GSK may limit generalization to diverse datasets. Al-Fatlawi *et al.* [12] proposed a fraud detection model for banking systems using genetic algorithms for feature selection and classification. The Statlog (German Credit Data) dataset was used, and results showed improved precision (90.4%) and accuracy (91.03%) post-feature selection. However, limitations include overfitting risks with decision trees and potential inefficiency for evolving fraud scenarios. Dasari and Kaluri [13] presented a privacy-preserving federated learning (FL) framework (2P3FL) using FedAvg and FedProx algorithms within the flower framework. It employs the CreditCard and CICIDS datasets. Achieving 99.57% accuracy on the CreditCard dataset, limitations include data quality issues and challenges with model convergence. The approach improves privacy and performance in distributed financial systems. Hussain *et al.* [14] proposed an enhanced intelligent intrusion detection system (NIDS) for e-commerce using an extended backward oracle matching (BOM) algorithm. Tested on NSL-KDD datasets and real traffic scenarios, it achieved a 5.17% higher detection rate and 0.22% fewer false alarms. Despite improved packet analysis, limitations include high packet drop rates under heavy traffic. Uddin *et al.* [15] introduced a dual-tier adaptive IDS using one-class classifiers (OCC) with semi-supervised learning and clustering (usfAD and DBSCAN). Tested on 10 datasets (e.g., NSL-KDD, UNSW-NB15), it identifies known/unknown attacks, overcoming zero-day detection challenges. Results showed accuracy improvements after retraining. Limitations include high resource requirements for clustering and scalability concerns. Vamsikrishna *et al.* [16] implemented a hierarchical anomaly IDS using artificial neural networks (ANN) in cloud computing environment. It processes high-traffic data efficiently, achieving superior accuracy (98.5%) compared to decision trees (91.3%). However, ANN's reliance on pre-selected features limit adaptability. Evaluation used datasets with mixed normal and malicious traffic to enhance robustness.

Building on these insights, this study employs the CICIDS2017 and CSECICIDS2018 datasets, which replicate realistic network traffic scenarios and encompass diverse attack types, including DoS, DDoS, Botnet, and Web-based threats [17], [18]. These datasets are well-suited for training IDS and fortifying banking sector security. Key preprocessing steps-such as data integration, encoding, and standardization-are applied to ensure compatibility and boost model efficiency. To address the high-dimensional characteristics of the data by utilizing a basic autoencoder for dimensionality reduction. The autoencoder preserves essential features while reducing input dimensions, thereby lowering computational overhead. To detect attacks, the study implements DBSCAN, a density-based clustering algorithm. DBSCAN stands out from traditional techniques by identifying clusters of arbitrary shapes and recognizing noise points, enabling it to effectively detect complex attack patterns without requiring prior knowledge of the cluster count. The proposed framework unifies these methods into a comprehensive solution, providing an adaptive and effective strategy for securing banking systems. This study advances the field of cybersecurity by showcasing how the integration of autoencoders and DBSCAN improves the detection of both established and emerging threats in the banking domain.

2. RESEARCH METHOD

The proposed method, adaptive IDS, is designed to detect intrusions in banking domain using cloud datasets and classify them into multiple classes, as shown in the architecture in Figure 1. This architecture includes dataset preprocessing, dimensionality reduction, and multiclass attack detection.

2.1. Datasets

The CICIDS2017 [19] and CSECICIDS2018 [20] datasets, developed by the Canadian Institute for Cybersecurity, serve as standard benchmarks for intrusion detection research. They replicate realistic network traffic by incorporating both benign and malicious activities [21]. The datasets feature a diverse range of attack scenarios, such as DoS, DDoS, Botnet, Infiltration, and Web-based threats, categorized into 14 distinct attack types. The datasets include multi-day network traffic logs stored in CSV format, with each file representing unique traffic characteristics. The CICIDS2017 dataset contains 83 features, while the CSECICIDS2018 dataset includes 80 features, covering details such as source and destination IPs, protocols, packet sizes, and temporal attributes. Their extensive representation of contemporary threats and diverse traffic patterns makes them well-suited for assessing intrusion detection methods, particularly in the banking sector.

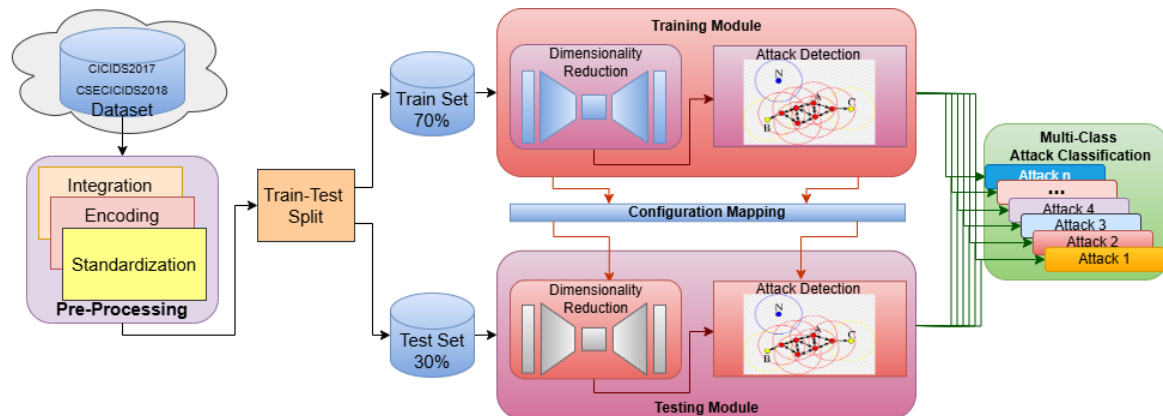


Figure 1. Architecture of proposed adaptive IDS

2.2. Preprocessing flow for intrusion detection in banking

The preprocessing stage is critical to structuring and preparing the CICIDS2017 and CSECICIDS2018 datasets for effective intrusion detection in the banking sector. This process involves key steps such as data integration, encoding, standardization, and splitting. Both datasets comprise multiple files, each capturing network traffic data collected across different days. This step involves combining all individual files into a unified dataset to ensure data consistency and completeness for analysis. The consolidated dataset often contains categorical attributes, which are converted into numerical formats. One-hot encoding is utilized to transform these categorical variables into binary vector representations. This approach preserves the distinct values of categorical attributes without implying any ordinal relationships. To maintain consistency in data scaling and improve the efficiency of machine learning models, standardization is performed. The Z-score standardization technique is applied, adjusting each feature to have a mean of 0 and a standard deviation of 1. This process removes bias stemming from variations in feature magnitudes, ensuring equal contribution from all features to the model. After preprocessing, the dataset is split into two subsets: 70% is allocated for training, and the remaining 30% is reserved for testing. This approach allows the model to identify patterns during training and evaluate its performance on unseen data in the testing phase. These preprocessing steps convert the raw dataset into a format suitable for dimensionality reduction, attack detection, and multi-class classification. Such a structured preprocessing workflow ensures dependable and accurate results for intrusion detection in the banking sector.

2.3. Dimensionality reduction process flow for intrusion detection

Dimensionality reduction plays a vital role in optimizing the CICIDS2017 and CSECICIDS2018 datasets for effective intrusion detection in the banking sector. This step simplifies the high-dimensional data into a more compact representation, preserving critical information. The preprocessed dataset, containing 82 and 79 numerical features respectively, is prepared for this reduction process. To maintain a consistent scale across variables, features are standardized using Z-score normalization. Dimensionality reduction is achieved using a basic autoencoder (bAE), an unsupervised neural network [22] designed to encode and compress data efficiently. The input layer corresponds to the 82 and 79 features present in the dataset. Through the encoder, the high-dimensional input is compressed into a latent representation consisting of 21 dimensions. The rectified linear unit (ReLU) activation function is applied in this layer, introducing non-linearity to facilitate the learning of intricate patterns effectively. This layer represents the compressed feature space, retaining the

most important information and eliminating unnecessary features. The bottleneck consists of 21 nodes, drastically reducing the dimensionality from 82 and 79. The decoder reconstructs the original input data from the 21-dimensional latent space, employing a sigmoid activation function to ensure smooth reconstruction. The autoencoder is trained using the Adam optimizer with a learning rate of 0.001, aiming to minimize reconstruction loss over 32 epochs with a batch size of 1024.

Throughout the training process, the autoencoder learns to compress the input data into the latent space and then reconstruct it to its original dimensions. The model is specifically trained on normal traffic from the training set (70% of the data), with reconstruction error serving as the loss function. After training, the encoder part of the autoencoder is used to extract the 21-dimensional latent features from the input data. These features create a condensed version of the original data, highlighting key patterns and relationships. By reducing the feature space from 82 to 21 dimensions, computational complexity is reduced, allowing for quicker processing in subsequent tasks. The autoencoder naturally eliminates noise, retaining only the most significant features. The compressed feature space streamlines the clustering process for detecting attacks, enhancing both accuracy and reliability. This dimensionality reduction not only prepares the dataset for the DBSCAN-based attack detection model but also ensures scalability and robustness, which are essential for intrusion detection in banking environments.

2.4. Attack detection using DBSCAN

The attack detection phase employs the density-based spatial clustering of applications with noise (DBSCAN) algorithm to detect and classify malicious activities within the reduced feature space. This step plays a crucial role in identifying intrusions and categorizing them into distinct attack types, facilitating effective mitigation, especially in the context of banking. The attack detection module utilizes the 21-dimensional feature set produced by the basic autoencoder (bAE) during the dimensionality reduction process, offering a compact and noise-free representation of the CICIDS2017 and CSECICIDS2018 datasets for clustering. DBSCAN, a density-based clustering algorithm [23], organizes data points based on their spatial closeness within the feature space, which makes it particularly effective for intrusion detection. Its ability to identify arbitrarily shaped clusters and manage noise allows it to isolate anomalies, potentially indicating new or previously unseen attack types. The essential parameters for DBSCAN are Epsilon (ϵ), set to 0.2, which determines the maximum allowable distance between points to be grouped together, and MinPts, set to 600, which defines the minimum number of points needed to form a dense region. The algorithm classifies points into three categories: core points, which have a minimum of 600 neighbors within a radius of 0.2; border points, which are within a core point's neighborhood but have fewer than 600 neighbors; and noise points, which do not belong to any cluster and are considered outliers or anomalies. Core and border points are clustered together based on density connectivity, with areas of high density forming separate clusters that correspond to different types of attacks.

DBSCAN clusters are then mapped to the attack categories defined in the CICIDS2017 and CSECICIDS2018 datasets. CICIDS2017 categories include DoS Hulk, PortScan, DDoS, DoS GoldenEye, FTP-Patator, SSH-Patator, DoS Slowloris, DoS SlowHTTPTest, Bot, Web Attack-Brute Force, and Web Attack-XSS. CSECICIDS2018 categories include HOIC, LOIC-UDP, and LOIC-HTTP in the DDoS family; Hulk, GoldenEye, SlowHTTPTest, and Slowloris in the DoS family; FTP and SSH in the Brute Force family; and Bot, Infiltration, and Web. The primary attack type for each cluster is assigned based on the majority label of the points it contains, whereas outliers (noise points) are examined further as potential novel or rare attack types that need more detailed investigation. The outcome of this phase is a multi-class classification of attacks that includes predefined categories. Each data point in the testing set is assigned a specific attack label based on its cluster, marked as noise, or classified as benign (normal) if it is identified as non-malicious traffic. Key outcomes include cluster assignments, where DBSCAN either groups points or designates them as noise; attack categories, where clusters are associated with corresponding attack types; and anomaly detection, where noise points are flagged as potential anomalies for further analysis.

DBSCAN efficiently manages noise by isolating new attacks that do not conform to existing clusters and can detect clusters of any shape, ensuring a precise representation of various attack patterns. As an unsupervised learning algorithm, DBSCAN does not rely on labeled data for training, making it ideal for real-world situations where attack labels may not be readily available. This is particularly important in the banking sector, where the sensitivity of the system allows for the early identification of anomalies in high-dimensional traffic data. DBSCAN's ability to handle noisy and imbalanced data is essential for dynamic banking networks that experience fluctuating traffic patterns. Its ability to classify into multiple categories divides attacks into 11 distinct groups, offering valuable insights for the implementation of targeted mitigation strategies. By transforming the reduced feature space into meaningful clusters that represent specific attack types, the DBSCAN-based process provides effective intrusion detection, customized to meet the security demands of the banking sector, thereby strengthening overall network resilience and security.

3. RESULTS AND DISCUSSION

This section provides a detailed analysis of the experimental results, emphasizing the performance evaluation of the proposed adaptive IDS model against conventional approaches. Key metrics such as precision, recall, F-measure, accuracy, and AUC-ROC curves are used to highlight the model's effectiveness in overcoming the challenges of intrusion detection.

The chart in Figure 2 depicts the mean squared error (MSE) trends over 40 epochs for CICIDS2017 and CSE-CICIDS2018 datasets during training and testing, using an autoencoder for dimensionality reduction. Initially, MSE decreases sharply, stabilizing around epoch 15. Both datasets show consistent convergence below the threshold (2.6986), indicating effective reconstruction of features. Training errors for CICIDS2017 and CSE-CICIDS2018 are marginally lower than test errors, reflecting the model's generalization capability. This suggests the autoencoder successfully reduces dimensions while preserving critical patterns.

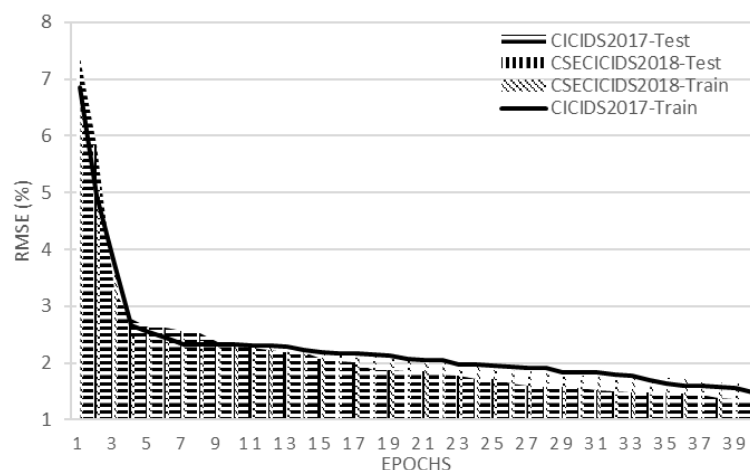


Figure 2. Trends of MSE during dimensionality reduction

The performance of the proposed model is evaluated against support vector machine (SVM), long short-term memory (LSTM), and K-Means on the CIC-IDS2017 dataset. Figure 3 illustrates a detailed comparison of their respective metrics. The proposed model significantly outperforms the others, achieving the highest precision (0.9948), recall (0.9907), F-measure (0.9927), and accuracy (0.9886). While SVM, LSTM, and K-Means show comparable results with slight variations, their metrics are consistently lower. This highlights the proposed model's superior ability to detect and classify malicious activities accurately, demonstrating its effectiveness for intrusion detection.

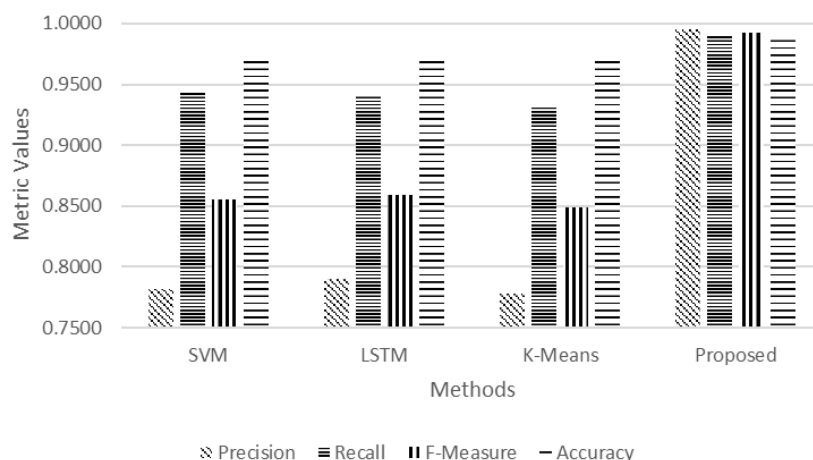


Figure 3. Performance evaluation of the proposed model on the CIC-IDS2017 dataset

Table 1 shows the performance of the proposed model compared to SVM, LSTM, and K-Means on the CSE-CIC-IDS2018 dataset. The proposed model achieves the highest precision (0.9966), recall (0.9901), F-measure (0.9933), and accuracy (0.9888), significantly outperforming the other models. While SVM, LSTM, and K-Means exhibit reasonable results, their performance metrics are consistently lower, particularly in precision and recall. This highlights the proposed model's superior effectiveness and reliability in detecting and classifying malicious activities in the dataset.

Table 1. Performance evaluation of the proposed model on the CSE-CIC-IDS2018 dataset

Measures	SVM [24]	LSTM [25]	K-Means [26]	Proposed
Precision	0.8304	0.8270	0.8120	0.9966
Recall	0.9287	0.9246	0.8995	0.9901
F-Measure	0.8768	0.8731	0.8535	0.9933
Accuracy	0.9689	0.9679	0.9632	0.9888

Figure 4 presents the AUC-ROC curves for the DBSCAN-based IDS applied to CICIDS2017 and CSE-CICIDS2018 datasets. The curves show a high true positive rate (TPR) for low false positive rates (FPR), indicating strong classification performance. Both datasets achieve near-perfect AUC scores, reflecting the model's ability to distinguish between benign and malicious traffic effectively. The overlap of curves suggests consistent performance across datasets. The dashed diagonal line represents random guessing, which the model significantly outperforms.

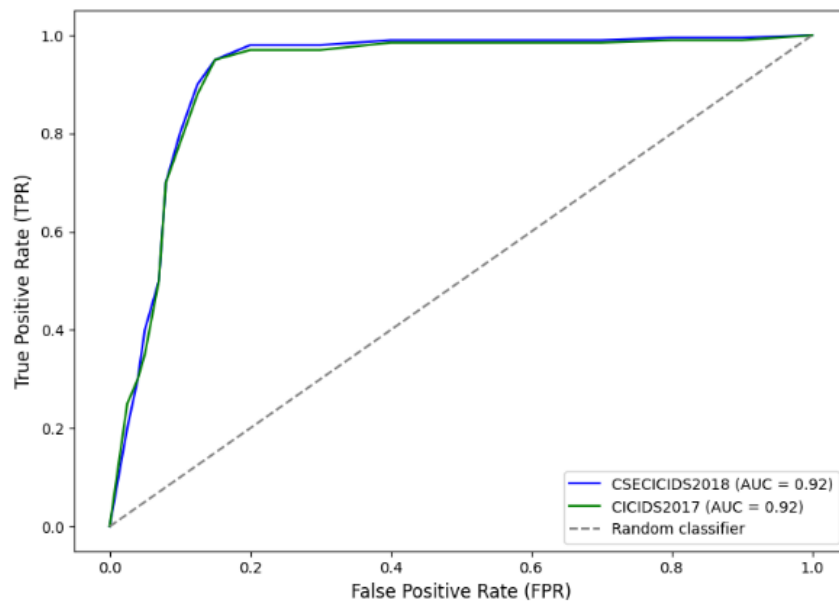


Figure 4. AUC-ROC curves for the adaptive IDS across datasets

Figure 5 shows the accuracy comparison of the proposed adaptive IDS model with SVM, LSTM, and K-Means across the CSE-CIC-IDS2018 and CIC-IDS2017 datasets. The proposed model consistently outperforms existing methods, achieving an accuracy of 0.9888 on CSECICIDS2018 and 0.9886 on CICIDS2017. While SVM, LSTM, and K-Means exhibit competitive accuracies (ranging between 0.9632 and 0.9731), the proposed model's superior performance highlights its robustness and reliability for intrusion detection across diverse datasets. This emphasizes its adaptability and effectiveness in addressing complex threat patterns.

The proposed adaptive IDS model demonstrates remarkable performance across multiple benchmarks, consistently surpassing traditional methods like SVM, LSTM, and K-Means in both CICIDS2017 and CSECICIDS2018 datasets. The autoencoder's efficient dimensionality reduction is evident from the MSE trends, where convergence below the set threshold affirms the model's ability to retain critical feature patterns. The slightly lower training errors compared to testing errors reflect the model's excellent

generalization capability. Performance metrics further validate the model's superiority. With the highest precision, recall, F-measure, and accuracy across both datasets, the proposed approach excels in identifying and classifying malicious activities. Particularly, its near-perfect AUC-ROC curves underscore its effectiveness in distinguishing between benign and malicious traffic. Moreover, the model's robustness across datasets highlights its adaptability to diverse traffic patterns and evolving threats. Overall, these results establish the proposed model as a reliable and advanced IDS, tailored to handle complex cyber-threat scenarios effectively.

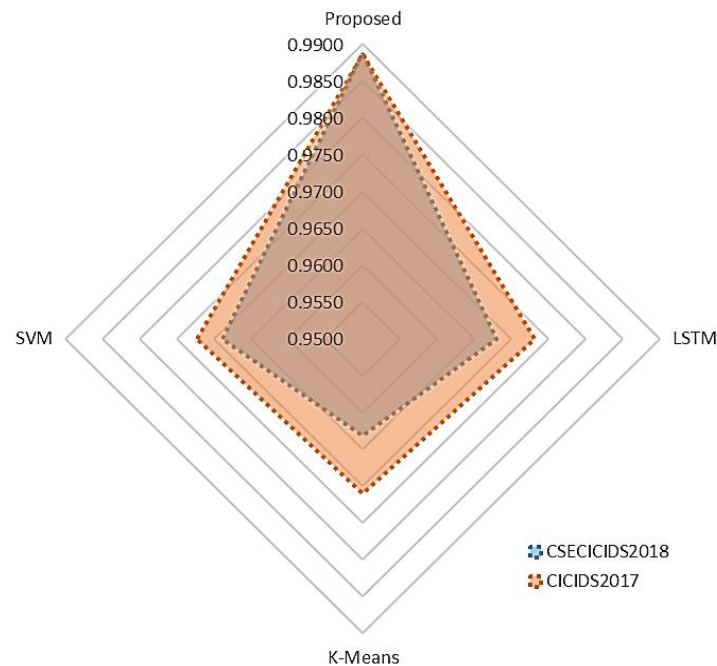


Figure 5. Accuracy comparison of the proposed adaptive IDS model with benchmarked models

4. CONCLUSION

This study presents a robust and adaptive intrusion detection framework specifically designed to enhance the security of banking systems. Leveraging the CICIDS2017 and CSECICIDS2018 datasets, the proposed adaptive IDS methodology integrates preprocessing, dimensionality reduction using a basic autoencoder, and attack detection with DBSCAN clustering to address key challenges such as high-dimensional data, imbalanced classes, and evolving cyber threats. The framework effectively classifies malicious activities into 11 distinct attack categories while maintaining computational efficiency. The evaluation demonstrates the framework's outstanding performance, achieving precision, recall, F1-score, and accuracy exceeding 98%. A comparative analysis with traditional methods, including SVM, LSTM, and K-means, confirms the superiority of the proposed approach in terms of detection accuracy and adaptability to complex attack patterns. By offering real-time detection capabilities and actionable insights, this framework provides a scalable solution for securing banking processes. Future work will focus on enhancing the adaptability of the framework to evolving attack vectors and integrating real-time threat intelligence for dynamic mitigation strategies. This research underscores the potential of combining machine learning techniques with clustering algorithms to strengthen cybersecurity in critical domains such as banking.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sathiyaseelan	✓			✓		✓	✓			✓	✓	✓		
Periyasamy														
Anubhav Kumar		✓				✓	✓			✓	✓		✓	
Karupusamy	✓				✓		✓			✓		✓		
Muthulakshmi														
Thenmozhi Elumalai	✓		✓		✓			✓	✓			✓		
Prabu Kaliyaperumal	✓	✓	✓	✓	✓			✓	✓	✓			✓	
Rajakumar Perumal		✓		✓		✓		✓	✓		✓			

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [P.K], upon reasonable request.





REFERENCES

- [1] E. Gülsleriler, B. Özgen, and Ş. Bahtiyar, "Malicious domain detection with machine learning for financial systems," in *2024 7th International Balkan Conference on Communications and Networking (BalkanCom)*, 2024, pp. 200–205, doi: 10.1109/BalkanCom61808.2024.10557171.
- [2] U. Islam *et al.*, "Detection of distributed denial of service (DDoS) attacks in IoT based monitoring system of banking sector using machine learning models," *Sustainability (Switzerland)*, vol. 14, no. 14, Jul. 2022, doi: 10.3390/su14148374.
- [3] Y. Gong, M. Zhu, S. Huo, Y. Xiang, and H. Yu, "Utilizing deep learning for enhancing network resilience in finance," in *2024 7th International Conference on Advanced Algorithms and Control Engineering (ICAACE)*, 2024, pp. 987–991, doi: 10.1109/ICAACE61206.2024.10549542.
- [4] D. O. Ogundipe, "Conceptualizing cloud computing in financial services: opportunities and challenges in Africa-US contexts," *Computer Science and IT Research Journal*, vol. 5, no. 4, pp. 757–767, Apr. 2024, doi: 10.51594/csitrj.v5i4.1020.
- [5] G. Muhammad, M. S. Hossain, and S. Garg, "Stacked autoencoder-based intrusion detection system to combat financial fraudulent," *IEEE Internet Things Journal*, vol. 10, no. 3, pp. 2071–2078, 2023, doi: 10.1109/JIOT.2020.3041184.
- [6] H. An *et al.*, "Finsformer: a novel approach to detecting financial attacks using transformer and cluster-attention," *Applied Sciences*, vol. 14, no. 1, p. 460, Jan. 2024, doi: 10.3390/app14010460.
- [7] M. Thankappan, N. Narayanan, M. S. Sanaj, A. Manoj, A. P. Menon, and M. G. Krishna, "Machine learning and deep learning architectures for intrusion detection system (IDS): A survey," in *2024 1st International Conference on Trends in Engineering Systems and Technologies (ICTEST)*, 2024, pp. 1–6, doi: 10.1109/ICTEST60614.2024.10576052.
- [8] M. J. Abudin, S. Thokchom, R. T. Naayagi, and G. Panda, "Detecting false data injection attacks using machine learning-based approaches for smart grid networks," *Applied Sciences (Switzerland)*, vol. 14, no. 11, Jun. 2024, doi: 10.3390/app14114764.
- [9] A. Kumar, R. Radhakrishnan, M. Sumithra, P. Kaliyaperumal, B. Balusamy, and F. Benedetto, "A scalable hybrid autoencoder-extreme learning machine framework for adaptive intrusion detection in high-dimensional networks," *Future Internet*, vol. 17, no. 5, p. 221, May 2025, doi: 10.3390/fi17050221.
- [10] M. Ozkan-Okay, R. Samet, Ö. Aslan, S. Kosunalp, T. Iliev, and I. Stoyanov, "A novel feature selection approach to classify intrusion attacks in network communications," *Applied Sciences (Switzerland)*, vol. 13, no. 19, Oct. 2023, doi: 10.3390/app131911067.
- [11] H. Q. Ghani and W. L. Al-Yaseen, "Two-step data clustering for improved intrusion detection system using CICIoT2023 dataset," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 9, Sep. 2024, doi: 10.1016/j.prime.2024.100673.
- [12] A. Al-Fatlawi, A. A. T. Al-Khazaali, and S. H. Hasan, "AI-based model for fraud detection in bank systems," *Fusion: Practice and Applications*, vol. 14, no. 1, pp. 19–27, 2024, doi: 10.54216/FPA.140102.
- [13] S. Dasari and R. Kaluri, "2P3FL: A novel approach for privacy preserving in financial sectors using flower federated learning," *CMES - Computer Modeling in Engineering and Sciences*, vol. 140, no. 2, pp. 2035–2051, 2024, doi: 10.32604/cmcs.2024.049152.
- [14] A. Hussain, K. N. Qureshi, K. Javeed, and M. Alhussein, "An enhanced intelligent intrusion detection system to secure e-commerce communication systems," *Computer Systems Science and Engineering*, vol. 47, no. 2, pp. 2513–2528, 2023, doi: 10.32604/csse.2023.040305.
- [15] Md. A. Uddin, S. Aryal, M. R. Bouadjene, M. Al-Hawawreh, and Md. A. Talukder, "A dual-tier adaptive one-class classification IDS for emerging cyberthreats," *Computer Communications*, vol. 229, p. 108006, 2025, doi: 10.1016/j.comcom.2024.108006.
- [16] M. Vamsikrishna *et al.*, "Cloud computing environment based hierarchical anomaly intrusion detection system using artificial neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 15, no. 1, pp. 1209–1217, Feb. 2025, doi: 10.11591/ijece.v15i1.pp1209-1217.





- [17] G. C. Amaizu, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "Investigating network intrusion detection datasets using machine learning," in *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, Oct. 2020, pp. 1325–1328, doi: 10.1109/ICTC49870.2020.9289329.
- [18] G. Singh and N. Khare, "A survey of intrusion detection from the perspective of intrusion datasets and machine learning techniques," *International Journal of Computers and Applications*, vol. 44, no. 7, pp. 659–669, 2022, doi: 10.1080/1206212X.2021.1885150.
- [19] Canadian Institute for Cybersecurity, "Intrusion detection evaluation dataset (CIC-IDS2017)." Accessed: Jan. 13, 2025. [Online]. Available: <https://www.unb.ca/cic/datasets/ids-2017.html>
- [20] Canadian Institute for Cybersecurity, "IPS/IDS dataset on AWS (CSE-CIC-IDS2018)." Accessed: Jan. 13, 2025. [Online]. Available: <https://registry.opendata.aws/cse-cic-ids2018/>
- [21] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, "Toward generating a new intrusion detection dataset and intrusion traffic characterization," in *ICISSP 2018 - Proceedings of the 4th International Conference on Information Systems Security and Privacy*, SciTePress, 2018, pp. 108–116, doi: 10.5220/0006639801080116.
- [22] K. Prabu, P. Sudhakar, M. Periyasamy, and A. Alagarsamy, "Harnessing DBSCAN and auto-encoder for hyper intrusion detection in cloud computing," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 13, no. 5, pp. 3345–3354, Oct. 2024, doi: 10.11591/eei.v13i5.8135.
- [23] D. H. Mustafa and I. M. Husien, "Adaptive DBSCAN with grey wolf optimizer for botnet detection," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 4, pp. 409–421, 2023, doi: 10.22266/ijies2023.0831.33.
- [24] M. A. Almaiah *et al.*, "Performance investigation of principal component analysis for intrusion detection system using different support vector machine kernels," *Electronics (Switzerland)*, vol. 11, no. 21, Nov. 2022, doi: 10.3390/electronics11213571.
- [25] H. K. Bella and S. Vasundra, "A novel framework based on extra tree regression classifier and grid search LSTM for intrusion detection in IoT and cloud environment," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 4, pp. 504–517, 2024, doi: 10.22266/IJIES2024.0831.39.
- [26] D. Dwivedi, A. Bhushan, A. K. Singh, and Snehlata, "Leveraging K-means clustering for enhanced detection of network traffic attacks," in *2024 3rd International conference on Power Electronics and IoT Applications in Renewable Energy and its Control (PARC)*, 2024, pp. 72–76, doi: 10.1109/PARC59193.2024.10486408.

BIOGRAPHIES OF AUTHORS






Sathiyaseelan Periyasamy     assistant professor in Department of Computer Science and Engineering at Chennai Institute of Technology, holds 12 years of teaching experience and is pursuing a Ph.D. in Computer Science and Engineering at Anna University. With an M.E. CSE from Anna University, specializing in networks, cloud computing, image processing, and machine learning. He can be contacted at email: sathiyaseelanp@citchennai.net.






Dr. Anubhav Kumar     professor in School of Computer Science and Engineering at Galgotias University, accumulating 24 years of teaching experience. He holds a Ph.D. and with 9 patents, 6 book chapters, and 18 research papers published in esteemed international journals and conferences. His expertise spans machine learning, deep learning, data science, NLP, and big data. He can be contacted at email: dr.anubhavkumar@gmail.com.






Dr. Karupusamy Muthulakshmi     associate professor in the Department of Information Technology at Panimalar Engineering College. With 19 years of teaching experience, she holds Ph.D. from Anna University and has authored 6 patents, 6 book chapters, and 19 research papers in renowned international journals and conferences. Her areas of expertise include cloud computing, cyber security, networks, and machine learning. She can be contacted at email: muthulakshmi_it@panimalar.ac.in.






Dr. Thenmozhi Elumalai    professor in the Department of Information Technology at Panimalar Engineering College. With 23 years of teaching experience, she holds Ph.D. and has authored 7 patents, 8 book chapters, and 24 research papers in renowned international journals and conferences. Her areas of expertise include cyber security, networks, and machine learning. She can be contacted at email: ethenmozhi22.pec@gmail.com.



Prabu Kaliyaperumal    assistant professor in School of Computer Science and Engineering at Galgotias University, has 17 years of teaching experience. Currently pursuing a Ph.D., he holds an M. Tech in CSE from SRM University and MBA from Anna University. He has published 4 patents, 2 book chapters and 17 research papers in international journals and conferences. His expertise includes cyber security, networks, cloud computing deep learning, and machine learning. He can be contacted at email: mega.prabu@gmail.com.



Rajakumar Perumal    assistant professor in Department of Computer Science and Engineering at Sharda University, holds 22 years of teaching experience and is pursuing a Ph.D. in Computer Science and Engineering at Shri Venkateshwara University. With an M.E. CSE from Anna University, he has published 4 patents and 8 research papers, specializing in networks, cloud computing, software engineering, and machine learning. He can be contacted at email: rajakumar.jcet@gmail.com.