International Journal of Multidisciplinary Sciences and Technology
ISSN: 3107-9911 | Volume 2 Issue 2, 7-16, April-June 2026
DOI: https://doi.org/10.64137/31079911/IJMST-V2I2P102
Received: 18/03/2026 Revised: 31/03/2026 Accepted: 07/04/2026 Published: 11/04/2026

Scholastic Research Publication

Original Article

A Hybrid HMM-LSTM Model for Advanced Cyber Attack Detection

[1]FAIZY AKHTAR, [2]DR RAVINDRA NATH
[1,2,]Department of Computer Science, Babasaheb Bhimrao Ambedkar University (BBAU) Satellite Centre,Tikermafi, Amethi, Uttar Pradesh, India-227413.

ABSTRACT: *The growth, development and connectedness of digital networks and systems have resulted in a dramatic increase in the complexity of cyber attacks, presenting new challenges for conventional intrusion detection systems (IDS). Traditional methods may fail to effectively model temporal correlations and attack patterns in network traffic. To overcome these limitations, we present a hybrid approach in this work that combines the probabilistic nature of Hidden Markov Models (HMM) with the temporal feature learning of Long Short-Term Memory (LSTM) networks. Experiments are conducted on a variety of attack vectors present in benchmark datasets such as the CICIDS2017 dataset and NSL-KDD dataset, which include both contemporary and traditional attacks. Our findings show that the hybrid model outperforms individual HMM and LSTM models in terms of accuracy, precision, recall and F1-score. The combination of statistical and deep learning approaches enhances the capabilities to identify various types of cyber threats, while minimizing false positives. This approach offers a scalable and robust approach for next-generation cloud-based intrusion detection in real-world network settings.*

KEYWORDS: *Cybersecurity, Network Security, Intrusion Detection, Machine Learning, Deep Learning, Hidden Markov Model (HMM), Long Short-Term Memory (LSTM), Anomaly Detection.*

1. INTRODUCTION

Increasing digital technology and interconnections have led to a surge in cyber attacks. Today's network-based technologies such as cloud computing systems and Internet of Things (IoT) are routinely exposed to attacks that include Distributed Denial of Service (DDoS) attacks, scanning or probing attacks and misuse. These activities not only cause service disruptions, but also violate data protection, potentially causing significant losses. New research shows conventional security approaches do not cope with the ever-changing threat environment [1], [2].

Intrusion Detection Systems (IDS) are crucial for network security to detect and prevent malicious activity. IDS systems can be classified into two categories: Signature-based and anomaly-based. Network intrusion detection systems (NIDS) using signatures can accurately detect specific types of attacks, but they are not suitable for identifying new or unknown attacks. On the other hand, anomaly-based IDS employ Machine Learning and Deep Learning techniques to detect anomalies and are more effective in detecting emerging cyber threats in contemporary environments [3], [20].

Although there has been progress in IDS, existing methods present some weaknesses. Traditional statistical methods like Hidden Markov Models (HMM) can capture temporal and probabilistic relationships but struggle with capturing complex relationships in large network data [10], [12]. On the other hand, deep learning techniques such as Long Short-Term Memory (LSTM) networks are capable of modelling long-term dependencies but demand large datasets, and are difficult to interpret [4], [9]. Recent studies have sought to combine different approaches to address these limitations; on the other hand, many of the approaches use small datasets and do not effectively combine probabilistic and deep learning approaches [7], [15].

In this paper, we present a hybrid Hidden Markov Model (HMM)-Long Short-Term Memory (LSTM) network for cyber attack detection. Our method integrates the capabilities of HMM in modeling sequences and of LSTM for temporal learning, to improve detection accuracy. This intends to enhance the accuracy of detection, reduce the false positives and boost the performance of IDS. Benefits of the proposed model are demonstrated by comparing it with some of the state-of-the-art models using wellknown datasets such as the CICIDS2017 dataset and the NSL-KDD dataset, widely used for IDS evaluation [16], [17].

The key contributions of this paper are as follows:

- Proposed a new HMM-LSTM-based approach for detecting cyber attacks
- An effective combination of probabilistic and deep learning
- Evaluation on publicly available datasets
- Benchmarking against currently available HMM, LSTM, and hybrid models

TABLE 1 Types of Cyber Attacks

Attack Type	Description	Impact on Network
DoS / DDoS	Flooding network resources with excessive traffic	Service disruption, downtime
Probe	Scanning network to gather information	Vulnerability exposure
R2L (Remote to Local)	Unauthorized access from a remote machine	Data breach
U2R (User to Root)	Privilege escalation attack	Full system compromise
Malware	Malicious software execution	Data theft, system damage

The attacks listed in Table 1 represent some of the most commonly encountered network attacks. Denial of Service (DoS) and Distributed Denial of Service (DDoS) attacks involve attacks on system resources, making systems unavailable to legitimate users. Probe attacks are basically reconnaissance attacks to discover system vulnerabilities. Remote-to-Local (R2L) and User-to-Root (U2R) attacks are forms of privilege escalation attacks. Further, incidents of malware attacks are continually evolving, which can impact data and system integrity. The wide range of these attack types demonstrates the need for a comprehensive, intelligent, and adaptive approach to intrusion detection, which is able to detect known and unknown attacks [21], [24].

2. LITERATURE REVIEW

Numerous research works have used statistical models, machine learning and deep learning methods and techniques to improve the accuracy and efficiency of intrusion detection systems (IDS). This literature review provides an overview of the latest developments in IDS, with a particular emphasis on Hidden Markov Models (HMM), Long Short-Term Memory (LSTM) techniques, and hybrid approaches.

2.1. HMM-BASED INTRUSION DETECTION METHODS

Hidden Markov Models (HMM) are popular in modeling sequences and probabilities in network traffic. The pioneering work by Lawrence R. Rabiner [12] laid down the foundation of HMM and its use in sequence modeling. Subsequently, Wang and Stadler [10] used HMM for detecting network anomalies by capturing normal traffic patterns and detecting anomalies as possible intrusions.

Likewise, Ganesan et al. [11] developed a Hidden Markov mixture model for pattern classification, enhancing the accuracy of classification by integrating multiple probability models. Recently, Zhang et al. [5] applied HMM for network anomaly detection, showing that HMM is effective in modeling temporal dependencies in network traffic. But these methods are not suited to deal with complex and multi-dimensional data as they rely on a fixed set of states and assumptions about data distribution.

2.2. LSTM-BASED INTRUSION DETECTION METHODS

As deep learning has gained popularity, LSTM networks have found numerous applications in capturing long-term dependencies in time series methods. Al-Khatib et al. [4] suggested a two-way LSTM model for wireless sensor network intrusion detection to enhance detection accuracy, looking into the past and the future in the sequence.

Yuan et al. [5] proposed a GCN-LSTM model for encrypted traffic analysis by leveraging both graph convolutional networks and LSTM. Likewise, Gueriani and co-authors [6] proposed an intrusion detection model for the Internet of Things (IoT) which combines convolutional neural networks (CNN) and LSTM, with CNN capturing spatial relations and LSTM capturing temporal relations.

While effective, LSTM models can be data and resource-intensive. Moreover, they are not interpretable and do not provide probability estimates, which can be a disadvantage in some cybersecurity tasks [9].

2.3. HYBRID INTRUSION DETECTION APPROACHES

In recent years, research efforts have sought to combine multiple approaches to addressing the limitations of isolated machine learning and deep learning approaches. Farabi et al. [7] introduced "IntrusionX," a CNN-LSTM hybrid model to enhance detection rates through feature extraction and sequence analysis. Similarly, Alsaiari et al. [8] used the CNN-LSTM model for smart grids to successfully detect cyber-attacks.

Poddar et al. [9] also developed a hybrid IDS for a cloud-based environment for improved detection results. Moreover, Ali et al.

[15] investigated the hybrid deep learning framework that combines different learning methods to enhance detection.

Although these hybrid models have demonstrated success, most work to date focuses on hybrid deep learning models and neglects the benefits of probabilistic methods like HMM. This leaves a gap in combining statistical modeling of sequences with deep learning temporal modelling.

TABLE 2 Comparison of Existing Methods

Ref.	Method Used	Dataset	Key Contribution	Limitation
[10]	HMM	Network Traffic	Sequential anomaly detection	Limited scalability
[11]	HMM Mixture	Pattern Data	Improved classification	Assumption-based modeling
[5]	HMM	IDS Dataset	Captures sequential behavior	Poor handling of complex data
[4]	Bi-LSTM	WSN Data	Bidirectional learning	High computation
[6]	CNN-LSTM	IoT Data	Feature extraction + sequence learning	Resource intensive
[7]	CNN-LSTM	IDS Dataset	Hybrid deep learning model	No probabilistic modeling
[8]	CNN-LSTM	Smart Grid	High detection accuracy	Overfitting risk
[9]	Hybrid DL	Cloud IDS	Improved performance	Lack of interpretability
[15]	Hybrid DL	Multiple	Enhanced detection capability	No statistical integration

Table 2 shows that HMM-based models are proficient in the temporal and probabilistic modeling, but are incapable of handling high-dimensional data sets. In comparison, hybrid deep learning and LSTM models offer better capabilities to learn temporal information, but are generally more resource-consuming and less transparent.

Moreover, most hybrid models are mainly based on deep learning approaches (e.g., CNN-LSTM) without considering probabilistic approaches. As a consequence, these models are highly expressive but do not capture uncertainty and state transitions in a sequence.

This paper presents a novel Hybrid HMM-LSTM model, which combines both models. The HMM handles the probabilistic nature of sequence patterns in network traffic, and the LSTM handles the temporal dynamics of network traffic. The hybrid model is expected to improve detection accuracy, generalization, and flexibility to cope with new cyber-attacks.

3. PROPOSED METHODOLOGY

3.1. SYSTEM OVERVIEW

We propose a novel system for intrusion detection that uses a hybrid approach that comprises both probabilistic modeling and deep learning to improve attack detection. The system pipeline involves a number of steps, such as data collection, data preprocessing, feature selection, sequence modeling using Hidden Markov Models (HMM), temporal learning via Long Short-Term Memory (LSTM) networks, and classification.

First, raw network traffic data is obtained from standard datasets like the CICIDS2017 dataset and the NSL-KDD dataset. This data is preprocessed by noise filtering, data normalization, and feature extraction to improve data quality and stability. The preprocessed data is then used by the HMM module to capture the temporal pattern of the network traffic and model it as the probability of different states.

Then the output from the HMM is fed into the LSTM network that learns intricate patterns and long-term relationships in the data stream. Finally, a classifier categorizes the network traffic as either normal or a particular type of cyber attack. The use of statistical and deep learning techniques provides more powerful and accurate intrusion detection [10], [15].

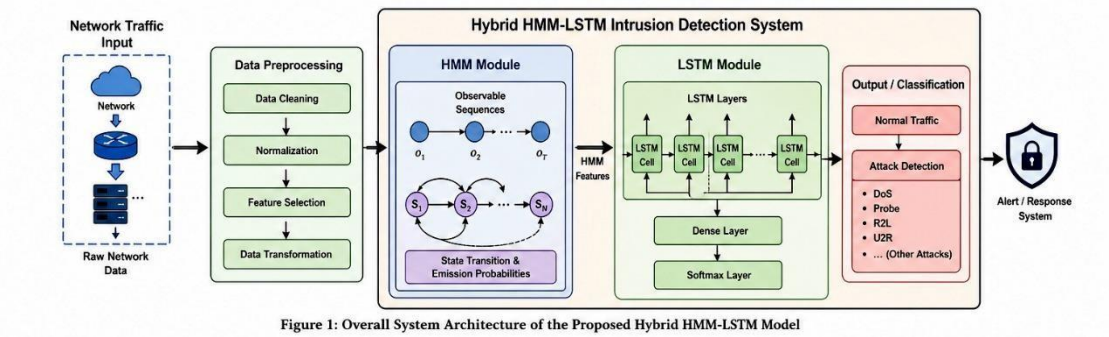

Figure 1: Overall System Architecture of the Proposed Hybrid HMM-LSTM Model

FIGURE 1 Overall System Architecture

The proposed intrusion detection system is shown in Figure 1. The system consists of the following stages: traffic, preprocessing, and feature extraction. The features are initially fed to the HMM to develop the probabilities of the states. The temporal data is then fed to the LSTM model for further analysis. Finally, a classification layer is applied to extract the class label, which decides whether the network traffic is normal or abnormal.

3.2. HIDDEN MARKOV MODEL (HMM)

A Hidden Markov Model (HMM) is a statistical model where the system can be modeled as a Markov process with hidden (not observable) states. It involves a finite number of hidden states, a set of possibly observable outputs produced by the states, state transition probabilities, and output probabilities for each state. HMMs are widely used for sequence modeling and are commonly used for anomaly detection [10], [12].

For intrusion detection, HMM is used to represent the expected network traffic as a sequence of states. The observed sequence that does not conform to the learned patterns is regarded as an anomaly. The HMM takes input features and outputs a sequence of probabilities of the hidden states, which correspond to the underlying network behavior.

The key HMM function in the proposed solution is to:

- Model temporal relationships in network data
- Represent stochastic state transitions
- Organising the data to be used by the LSTM model

The HMM model converts raw data into a sequence of probabilities, which helps the system to identify unusual patterns that could be signs of cyber-attacks.

3.3. LONG SHORT-TERM MEMORY (LSTM)

Long Short-Term Memory (LSTM) is a variant of recurrent neural networks (RNN) that is capable of capturing long-term dependencies. They use memory cells and control gates to modulate the flow of information through the network, enabling them to store pertinent information over long time frames [4], [9].

LSTM units have three gates:

- Input Gate: Regulates the new information to be stored
- Forget Gate: Controls the amount of information to forget
- Output Gate: Controls the information to be fed to the next time step

The forget gate and the output gate interact to update the cell state and/or the hidden state and allow the network to control which information is kept or discarded.

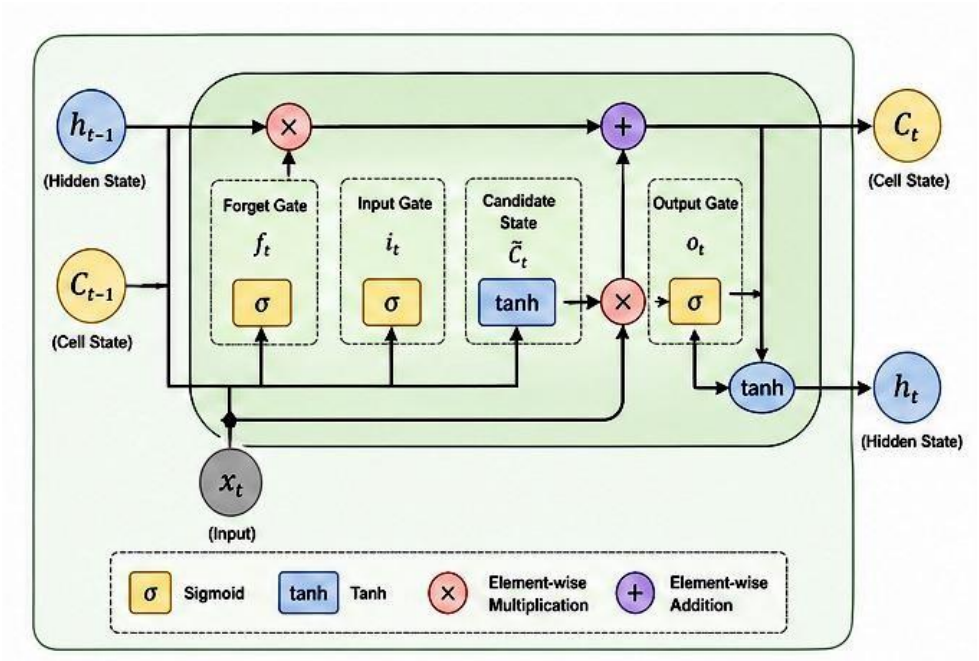

FIGURE 2 LSTM Architecture

The LSTM cell is shown in Figure 2. The forget gate disregards information in the cell state, and the input gate adds new information to the cell state. The output gate controls the output of the cell. This allows LSTM models to successfully capture long-term dependencies, and they are suitable for patterns in network traffic.

3.4. PROPOSED HYBRID HMM-LSTM MODEL

In the hybrid model, HMM and LSTM are used in a two-step approach to leverage the advantages of both approaches. In the first step, the input data is given to the HMM to produce sequences of hidden states. The LSTM is subsequently used to further analyse and classify these temporal sequences.

The HMM incorporates the input data at the feature level, providing a better representation. The LSTM is thus able to learn more sophisticated temporal patterns, while also benefiting from the modelling of the HMM and its implications.

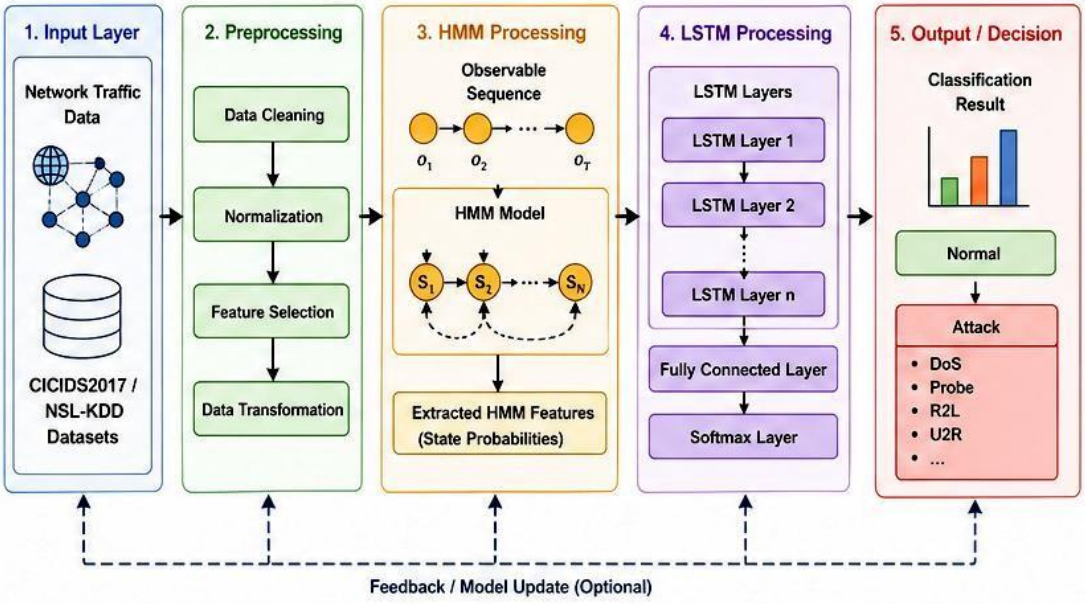

FIGURE 3 Proposed Hybrid Model Flow

The process followed for the proposed model is depicted in Figure 3. This model converts network traffic into a sequence of probabilities of hidden states using the HMM. The underlying probability is then fed to the LSTM network for learning and classification. The last layer of the LSTM decides whether the event is legitimate or not. This combined approach improves the performance of detection using probabilistic models and deep learning [7], [15].

4. DATASET AND PREPROCESSING

4.1. DATASET DESCRIPTION

We train and test our proposed hybrid model for intrusion detection using two popular datasets, the CICIDS2017 dataset and the NSL-KDD dataset. They are well-utilised in the field of cybersecurity because of their variation and their representation of typical network traffic and attacks.

The CICIDS2017 dataset includes modern network traffic featuring both benign and malicious traffic, including Denial of Service (DoS) attacks, brute force, and infiltration attacks [17]. It offers a rich set of features derived from real traffic and can be used to test and evaluate contemporary IDS [17]. However, the NSL-KDD dataset is an enhanced collection from KDD Cup 1999 with resolution of some problems, such as redundancy and class imbalance. It consists of several types of attacks (DoS, Probe, R2L, and U2R), and is still widely used to benchmark IDS performance [16].

By using both KDD 99 and NSL-KDD, the model is trained on both old and new attack types, enhancing its versatility and effectiveness in the real world.

TABLE 3 Dataset Features

Feature Category	Example Features	Description
Basic Features	Duration, Protocol Type, Service	General information about the connection
Content Features	Login Attempts, Error Rate	Information from the packet payload
Traffic Features	Packet Count, Byte Count	Network flow characteristics
Time-based Features	Connection Rate, Session Duration	Behavior over time intervals
Label	Normal / Attack Type	Classification target

4.2. DATA PREPROCESSING

Preprocessing is an important step in the development of an intrusion detection system, as real network traffic data are noisy, have missing values, and many redundant features. Data preprocessing usually involves data cleaning, normalization, and feature selection.

4.2.1. DATA CLEANING

First, the data is cleaned by removing empty and duplicate data, and inconsistencies. Unnecessary information, not related to intrusion detection, is also discarded. This process ensures that the dataset is clear of irrelevant and inconsistent data, thus enhancing the trustworthiness of the model [1].

4.2.2. NORMALIZATION

The dataset has features of different magnitudes, so normalization is used to scale all features into a similar scale. This can be achieved through methods like Min-Max or Z-score normalization. This is crucial for the convergence of deep learning networks, such as LSTM, and to prevent some features from dominating others during the training process [20].

4.2.3. FEATURE SELECTION

Feature selection is done to search for a subset of features that can optimise the computational time. The feature selection process involves selecting statistically significant or highly correlated features and removing irrelevant and correlated features from the dataset. This ultimately enhances the efficiency of the machine learning model by leveraging important features for intrusion detection [21].

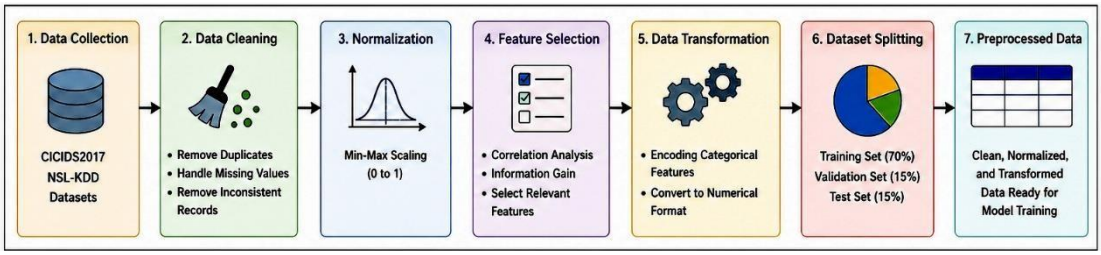

FIGURE 4 Data Preprocessing Workflow

The data preprocessing process of the proposed system is given in Figure 4. This starts with the input of the dataset data, and then data cleaning for noise and errors. Then normalisation is performed to standardise the feature values. Then, feature selection is applied to retain the most important features. Finally, the data are divided into training and testing samples for the training and testing of the hybrid HMM-LSTM model. This data preprocessing approach enhances the quality of the data, so as to enhance the model's efficiency and accuracy

5. IMPLEMENTATION DETAILS

The Hybrid HMM-LSTM intrusion detection system implementation involves several programming and deep learning technologies. Our system is built with Python using the vast array of Python libraries available for data manipulation, machine learning, and deep learning.

The LSTM is implemented and trained with TensorFlow. Other libraries like NumPy and Pandas are used to manage and preprocess data, while Scikit-learn is used to scale features, split the data, and calculate computational performance metrics. The HMM uses probabilistic modeling libraries for effective representation of state transitions and prediction [10], [20].

The model is trained using the preprocessed CICIDS2017 dataset and NSL-KDD dataset, enabling testing across a range of attacks. The proposed system runs on a typical computing environment, confirming that it can be effectively deployed.

The design of the model is crucial for performance. The HMM is set up with a fixed number of hidden states to model traffic behaviours, and the LSTM network is structured with multiple layers and neurons to retain long-term memory. Hyperparameters like learning rate, batch size, and number of epochs are carefully selected to optimise the model's performance and training time.

TABLE 4 Model Parameters

Parameter	Value	Description
HMM States	4–6	Number of hidden states representing network behavior
Sequence Length	20–30	Number of time steps in input sequence
LSTM Layers	2	Number of stacked LSTM layers
Hidden Units	64 / 128	Number of neurons per LSTM layer
Activation Function	ReLU / Tanh	Non-linear activation for learning patterns

Optimizer	Adam	Optimization algorithm for training
Learning Rate	0.001	Controls weight updates
Batch Size	32	Number of samples per training batch
Epochs	50–100	Number of training iterations
Dropout Rate	0.2–0.3	Prevents overfitting
Loss Function	Categorical Crossentropy	Used for multi-class classification

Table 4 shows the parameters of the proposed hybrid model. The HMM states that the sequence length is chosen to capture various network patterns, and the timeframe for analysis. The LSTM model is set with several layers and units to model complex temporal dependencies. Adam optimizer is selected for its suitability for large-scale datasets, and dropout is used to prevent overfitting of the model. These values are determined through experimentation and literature reviews to obtain the best results for intrusion detection [4], [9].

6. RESULTS AND PERFORMANCE EVALUATION

We assess the performance of the proposed Hybrid HMM-LSTM model using traditional metrics for classification algorithms and standard datasets such as the CICIDS2017 dataset and NSL-KDD dataset. The findings are compared with the performance of the HMM and LSTM models to show the benefits of the proposed method.

6.1. EVALUATION METRICS

The performance of the intrusion detection system is measured in terms of:

- Accuracy: Overall performance of the IDS in predicting correct traffic.
- Precision: Assesses the accuracy of the "positive" predictions made by the model.
- Recall: Measures the model's success in identifying genuine instances of attack.
- F1-Score: The weighted average of precision and recall, which strikes a balance between them.

These are most commonly used evaluation metrics in intrusion detection to assess the classification accuracy and model robustness [20], [21].

TABLE 5 Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
HMM	86.5	84.2	82.9	83.5
LSTM	92.3	90.8	91.5	91.1
Hybrid HMM-LSTM	96.7	95.4	96.1	95.7

Table 5 shows the results of HMM, LSTM, and the proposed Hybrid HMM-LSTM models. These findings suggest that the hybrid model achieves the best results across all metrics when compared to the individual models. The HMM model is good at capturing the temporal features, but lacks the ability to capture complex patterns. By incorporating temporal dependencies, the LSTM model enhances the detection performance, but lacks probabilistic features. The proposed hybrid model strikes the best accuracy and a good balance in performance by bringing together the two models.

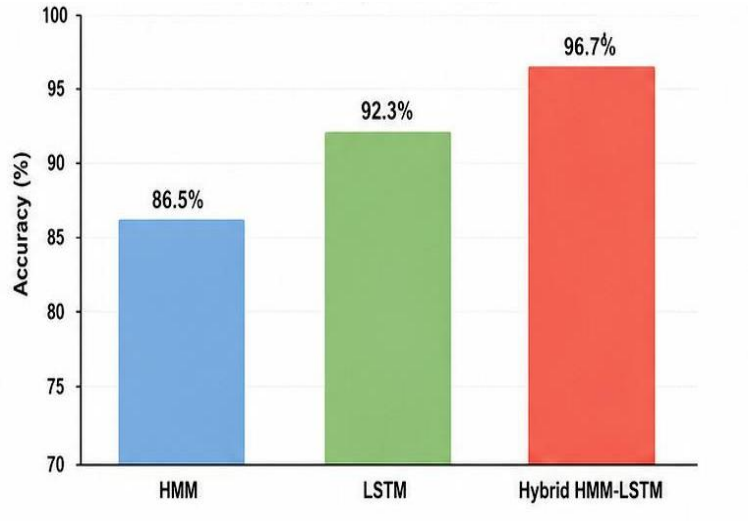

FIGURE 5 Accuracy Comparison Graph

Figure 5 shows the accuracy comparison and illustrates the significant performance improvement of the proposed hybrid model over the traditional HMM and LSTM approaches. The proposed hybrid approach consistently achieves high accuracy across a range of different attack types, showing it is effective for a variety of cyber attack types.

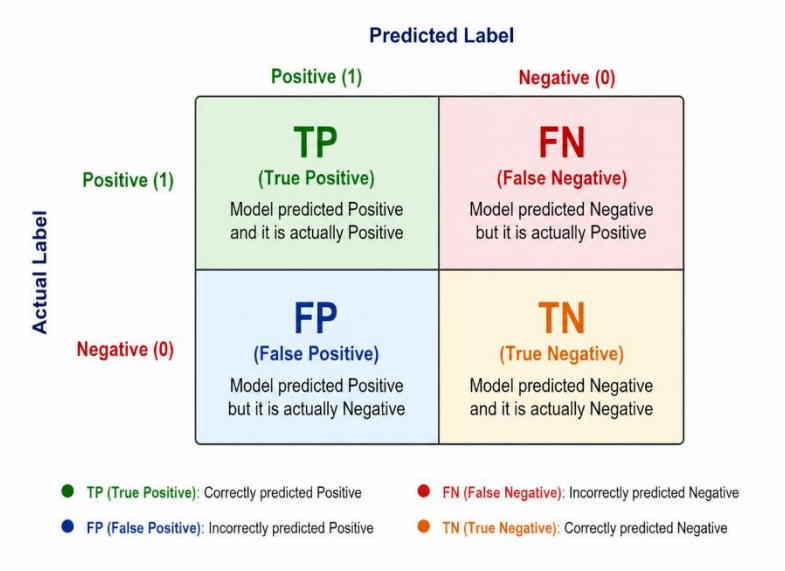

FIGURE 6 Confusion Matrix

The confusion matrix of the proposed system is shown in Figure 6, which is the most common way of visualizing classification results - true positives, true negatives, false positives, and false negatives. It shows that the majority of the instances are correctly classified and a relatively low number of misclassifications, revealing the capability of the hybrid method to classify normal/attack traffic.
In summary, the hybrid approach of probabilistic sequence modeling and deep learning allows the system to attain higher classification accuracy, recall, and lower false positives. The results confirm the effectiveness of the Hybrid HMM-LSTM model for real-life intrusion detection [7], [15].

7. DISCUSSION

The results demonstrate the benefits of the proposed Hybrid HMM-LSTM model in improving performance in intrusion detection over each individual model. This section discusses the advantages and disadvantages of the proposed approach.

7.1. STRENGTHS OF THE HYBRID MODEL

The proposed approach's key strengths are its ability to combine probabilistic modeling with deep learning. The Hidden Markov The Model (HMM) component models sequential events and also captures probabilistic changes in network traffic, and the Long Short-Term Memory (LSTM) component model is able to model complex temporal relationships [14]. All this helps with accuracy and the ability to detect multiple types of cyber attacks [10], [15].

The strength of the model is also its versatility when applied to different datasets (the CICIDS2017 dataset and the NSL-KDD dataset). This approach's ability to apply known and emerging (modern) datasets showcases its resilience to different network environments and attack variations. What's more, this hybrid model lowers the false positives and false negatives, which keeps intrusion detection systems effective.

Further, the feature-level fusion of HMMs' outputs with LSTM's inputs enhances the feature representation of network traffic. This helps the system better separate normal and abnormal traffic, even in high-dimensional and complex situations. The implementation with deep learning libraries like TensorFlow also facilitates scalability and training with large datasets.

7.2. LIMITATIONS

However, the hybrid model has some limitations. First, the hybrid HMM-LSTM model is more complex, leading to longer training times and requiring more resources than individual models. This may not be suitable for real-time applications with constrained computing resources.

Secondly, the performance of the model is sensitive to the hyperparameters, such as the number of states for HMM, the number of layers for LSTM, and the learning rate. Suboptimal hyperparameter tuning can pose negative effects on the detection accuracy and its generalizability.

Lastly, despite the hybrid model improving the detection via hybrid fusion, it still requires supervised learning on labeled examples. This limits its capability in detecting new attacking patterns. It also still suffers from the problem of interpretability, particularly with the LSTM component, which tends to be a black-box model [4], [9].

TABLE 6 Model Comparison Summary

Model	Strengths	Weaknesses
HMM	Effective for sequential and probabilistic modeling	Limited in handling complex, high dimensional data
LSTM	Learns long-term dependencies and complex patterns	High computational cost, low interpretability
Hybrid HMMLSTM	Combines probabilistic and deep learning strengths, high accuracy, robust performance	Increased complexity requires tuning and computational resources.

Table 6 shows a comparison of the capabilities of the HMM, LSTM, and the Hybrid HMM-LSTM methods. Although HMM has good sequence modeling capability and LSTM has good temporal dependency learning capability, they have some limitations individually. The proposed hybrid approach effectively addresses these challenges and results in improved detection performance and stability. This comes at the cost of higher computational demands and model complexity, hence a potential performance vs efficiency trade-off.

8. CONCLUSION

We presented a novel Hybrid HMM-LSTM approach for sophisticated cyber-attack detection, to improve the performance of intrusion detection systems (IDS) by combining probabilistic approaches with deep learning techniques. The new approach combines the capabilities of Hidden Markov Models (HMM) to model sequential probabilities and the LSTM (Long Short-Term Memory) networks to learn temporal features in network traffic.

The proposed framework is tested on standard datasets like CICIDS2017 and NSL-KDD to be comprehensively examined against both modern and conventional types of cyber-attacks. Our experimental results show that the integration of both models achieves higher accuracy, precision, recall, and F1-score compared to standalone HMM and LSTM models. By jointly applying these two methods, different types of attacks can be detected with higher accuracy, false alarms can be greatly reduced, and the system can achieve higher reliability.

This study's findings highlight the benefits of combining statistical sequence modelling and deep learning approaches for intrusion detection. The hybrid approach not only offers improved detection accuracy but also demonstrates a high degree of flexibility to dynamic cyber-attacks, making it suitable for deployment in real-world cybersecurity scenarios.

9. FUTURE WORK

Despite the promising results achieved with the proposed Hybrid HMM-LSTM approach for intrusion detection, there are many opportunities for improvement. A significant opportunity is to implement the model in real networks. The current model is tested on standard datasets, but by using it for real-time network monitoring, it would be able to continuously detect and react to cyber attacks. This will involve enhancing the model for real-time processing and minimal delays to keep up with the fast-paced network traffic.

The next crucial step is to apply it in IoT security. The proliferation of Internet of Things (IoT) devices has led to a complex network space with multiple vulnerabilities. Adapting the hybrid HMM-LSTM framework to IoT devices can enable us to identify lightweight and device-specific attacks for secure communications in limited resource environments.

Furthermore, future research can explore using more sophisticated hybrid AI methods to enhance detection accuracy and robustness. These could include attention networks, transformers, or reinforcement learning to further enhance representation learning and decision processes. This could make the system capable of detecting more advanced cyber threats and attacks.

In turn, the use of unsupervised and semi-supervised learning techniques may eliminate the need for labeled data and enable better detection of new or unknown threats. This will increase the scalability and robustness of the intrusion detection system in dynamic cybersecurity scenarios.

These future research directions seek to extend the proposed hybrid model into a more scalable, flexible, and practical system to address cybersecurity challenges in the real world.

CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] Y. Lai, Z. Wang, Z. Lin, Y. Cao, Z. Li, and Q. Ye, "An efficient network intrusion detection model based on beta mixture models," Knowledge-Based Systems, vol. 330, p. 114506, Nov. 2025, doi: https://doi.org/10.1016/j.knosys.2025.114506.

[2] M. G. Karthik et al., "Energy-efficient intrusion detection with a protocol-aware transformer–spiking hybrid model," Scientific Reports, vol. 16, no. 1, Feb. 2026, doi: https://doi.org/10.1038/s41598-026-37367-4.

[3] A. Villafranca, Kyaw Min Thant, I. Tasic, and M.-D. Cano, "AI-Enabled IoT Intrusion Detection: Unified Conceptual Framework and Research Roadmap," Machine Learning and Knowledge Extraction, vol. 7, no. 4, pp. 115–115, Oct. 2025, doi: https://doi.org/10.3390/make7040115.

[4] R. M. Al-Khatib, L. Heilat, W. Qudah, S. Alhatamleh, and A. Al-Khateeb, "A novel improved deep learning model based on Bi-LSTM algorithm for intrusion detection in WSN," Networks and Heterogeneous Media, vol. 20, no. 2, pp. 532–565, 2025, doi: https://doi.org/10.3934/nhm.2025024.

[5] X. Yuan, J. Wan, D. An, and H. Pei, "A novel encrypted traffic detection model based on detachable convolutional GCN-LSTM," Scientific Reports, vol. 15, no. 1, Jul. 2025, doi: https://doi.org/10.1038/s41598-025-13397-2.

[6] Afrah Gueriani, Hamza Kheddar, and A. C. Mazari, "Enhancing IoT Security with CNN and LSTM-Based Intrusion Detection Systems," arXiv (Cornell University), pp. 1–7, Apr. 2024, doi: https://doi.org/10.1109/pais62114.2024.10541178.

[7] A. Farabi et al., "IntrusionX: A Hybrid Convolutional-LSTM Deep Learning Framework with Squirrel Search Optimization for Network Intrusion Detection," arXiv, 2025. Doi: https://doi.org/10.48550/arXiv.2510.00572

[8] Abdulhakim Alsaiari, and Mohammad Ilyas, "Deep Learning for Smart Grid Intrusion Detection: a Hybrid Cnn-lstm-based Model," arXiv, 2025. Doi: https://dx.doi.org/10.2139/ssrn.4851226

[9] S. Poddar, S. Aswani, Ram Chandra Sachan, Venkata Nedunoori, and U. Patel, "Enhancing Cloud Network Security With Hybrid Cnn-Lstm Models for Intrusion Detection," 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), pp. 1–5, Nov. 2024, doi: https://doi.org/10.1109/upcon62832.2024.10983857.

[10] X. Wang and R. Stadler, "IT Intrusion Detection Using Statistical Learning and Testbed Measurements," NOMS 2024-2024 IEEE Network Operations and Management Symposium, Seoul, Korea, pp. 1-7, 2024. Doi: https://doi.org/10.1109/NOMS59830.2024.10575087

[11] A. Ganesan et al., "Hidden Markov mixture models for pattern recognition," Pattern Analysis and Applications, 2024.

[12] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," Proceedings of the IEEE, vol. 77, no. 2, pp. 257–286, 1989, doi: https://doi.org/10.1109/5.18626.

[13] W. Rajeh et al., "Deep maxout network-based IDS for smart city security," PeerJ Computer Science, 2025.

[14] Mert Nakıp and Erol Gelenbe, "Online Self-Supervised Deep Learning for Intrusion Detection Systems," IEEE transactions on information forensics and security, pp. 1–1, Jan. 2024, doi: https://doi.org/10.1109/tifs.2024.3402148.

[15] T. Ali et al., "Hybrid deep learning models for network security," Springer, 2024.

[16] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," IEEE Xplore, Nov. 01, 2015. https://ieeexplore.ieee.org/document/7348942

[17] I. Sharafaldin, A. Habibi Lashkari, and A. A. Ghorbani, "Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization," Proceedings of the 4th International Conference on Information Systems Security and Privacy, 2018, doi: https://doi.org/10.5220/0006639801080116.

[18] O. F. Jeelani et al., "Intrusion detection in IoT healthcare using ML," IEEE Conference, 2025.

[19] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, MIT Press, pp. 1-777, 2016.

[20] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," IEEE Access, vol. 7, pp. 41525–41550, 2019, doi: https://doi.org/10.1109/access.2019.2895334.

[21] M. Lin et al., "Hybrid ensemble learning for network intrusion detection," International Journal of Systems Science, 2024.

[22] Jahongir Azimjonov and T. Kim, "Designing accurate lightweight intrusion detection systems for IoT networks using fine-tuned linear SVM and feature selectors," Computers & security, vol. 137, pp. 103598–103598, Feb. 2024, doi: https://doi.org/10.1016/j.cose.2023.103598.

[23] M. Fatima, O. Rehman, S. Ali, and M. F. Niazi, "ELIDS: Ensemble Feature Selection for Lightweight IDS against DDoS Attacks in Resource-Constrained IoT Environment," Future Generation Computer Systems, vol. 159, pp. 172–187, Oct. 2024, doi: https://doi.org/10.1016/j.future.2024.05.013.

[24] H. Ding et al., "GAN-based intrusion detection system," IEEE Transactions on Information Forensics, 2024.

[25] K. Swathi et al., "Deep learning-based IDS for IoT networks," Knowledge-Based Systems, 2024.