A2S-AFLNet: An Adaptive Bat Optimized Two-Stage Attention Fused LSTM Networks for Attack-Resilient Intrusion Detection

Sagar Mankoti, Shukla Mondal, SK Hafizul Islam

Department of Computer Science and Engineering, Indian Institute of Information Technology Kalyani
{sagarmankoti,shuklamondal.cse,hafi786}@gmail.com

Abstract

With the increasing sophistication of cyber security risks, utilizing machine learning algorithms alongside intrusion detection systems has become crucial. Conventional approaches to detection intrusion on network traffic data come with limitations such as higher false positive rates, inability to adapt evolving attack patterns, and ineffective handling of large data volume. Deep neural networks such as long short-term memory (LSTMs) are good at understanding patterns in network data over time. Sometimes, they overlook aspects that can cause unnecessary calculations, which leads to less optimal detection results. We present an adaptive bat-optimized and two-stage LSTM network fused with attention (A2S-AFLNet) to address these issues. This method combines attention mechanisms and LSTM to improve feature selection while also making the learning process more efficient and adaptable for intrusion detection. Standard performance metrics were analyzed and compared with the recent machine learning and neural network-based IDS models using the UNSW-NB15 dataset to validate the robustness of the framework.

Introduction

We have seen a huge rise in internet users and connected devices in recent years. As a result, cybersecurity has become more important than ever. With the growing network traffic, detecting intrusions is now crucial to protecting sensitive data and maintaining the system's efficiency. An Intrusion Detection System (IDS) serves as an important component in maintaining cybersecurity by identifying and preventing potential vulnerabilities present in network traffic. Traditional rule-based IDS models struggle to effectively detect novel and complex attack patterns (Li, Li, and Li 2025). Machine learning and Artificial Neural Network (ANN) algorithms have emerged as a powerful alternative that utilizes data-driven techniques to enhance detection rates. However, these models often suffer from high computational complexity and degraded performance due to irrelevant or redundant features in high-dimensional network traffic data (Al-Haija and Droos 2025).

Network intrusion patterns are continuously evolving as the attackers disguise malicious activities within normal traffic to evade detection. Therefore, IDS models must be computationally efficient and adaptable to evolving network patterns. In this paper, we utilize adaptive Bat-based feature engineering as an effective feature selection mechanism to optimize IDS performance. Inspired by the echolocation behaviour of bats, this method balances exploration and exploitation to identify the most informative features from high-dimensional network datasets. The selected features are then forwarded to a deep learning-based IDS to improve detection efficiency and accuracy. This approach ensures that the most contributing network features are utilized to build the IDS for high detection rates while reducing overhead.

Traditional IDS often does not keep up with evolving attack patterns, requiring an adaptive approach to improve its effectiveness. The proposed IDS framework, A2S-AFLNet, includes a novel two-stage LSTM-based network incorporating attention to feature prioritization and the adaptive Bat-based optimization algorithm to enhance feature selection and efficiency. The bidirectional-LSTM (BiLSTM) layers capture temporal dependencies in sequential network data, and the attention mechanism highlights critical traffic patterns while improving detection accuracy. Extensive experiments are conducted utilizing the UNSW-NB15 dataset to validate the proposed approach, comparing its performance against recent machine learning and neural network-based IDS models.

The key contributions include:

- An adaptive bat is incorporated in the proposed IDS framework utilizing multi-objective fitness function and velocity adjustment to optimize network traffic features to reduce dimensionality while maintaining high classification accuracy.
- A two-stage LSTM-based architecture is designed to capture sequential dependencies in network traffic, whereas the first BiLSTM layer captures bidirectional temporal relationships while the second BiLSTM refines extracted features.
- Multi-head attention is integrated to focus on critical time steps and fused with residual connections to preserve initial traffic patterns, effectively detecting novel cyber threats.

IDS has evolved significantly with the integration of deep

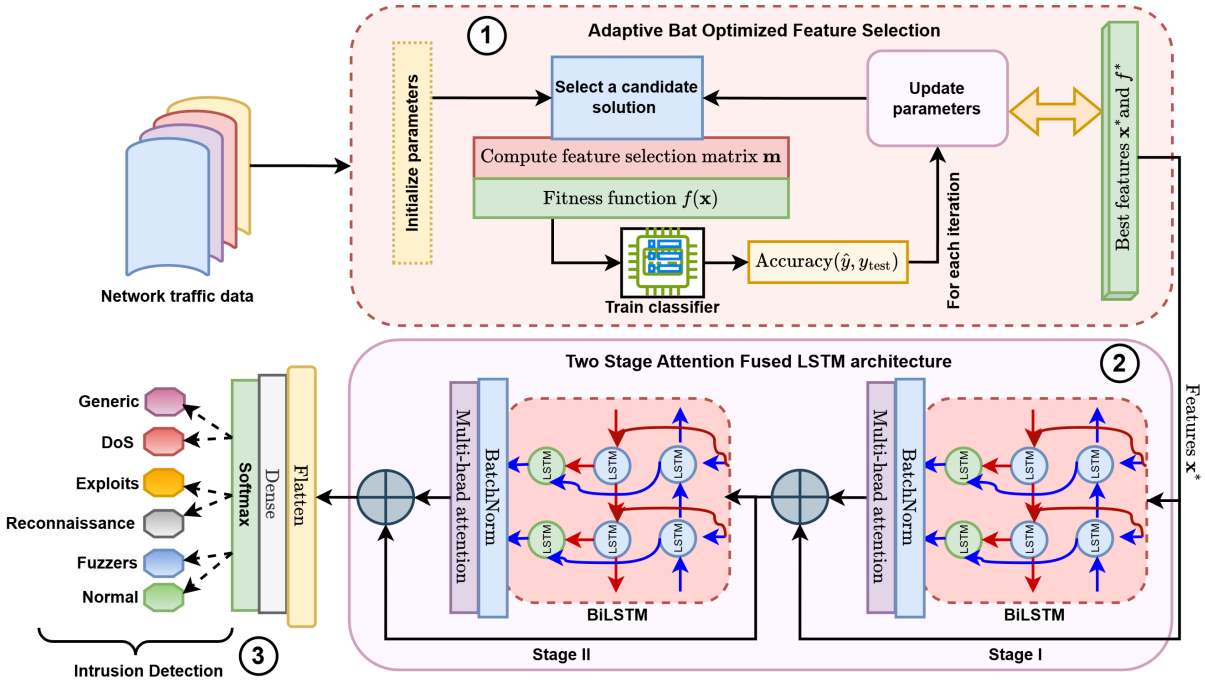

Figure 1: Overview of the proposed A2S-AFLNet to detect network intrusion

learning models, attention mechanisms, and bio-inspired optimization techniques. However, while recognizing temporal patterns effectively, conventional deep learning models, such as LSTMs, often struggle with irrelevant features and computational inefficiencies (Yu et al. 2022). Attention mechanisms help to focus on crucial features, reducing unnecessary computations (Wang et al. 2024a), while bio-inspired algorithms like bat (Yang 2010) optimize feature selection and parameter tuning (Xu, Qin, and Zhou 2022). The proposed A2S-AFLNet integrates these three approaches, LSTM, attention mechanisms, and adaptive bat optimization, to create a robust and adaptive intrusion detection framework.

Bat Optimized Feature Selection in IDS

The feature optimization is an important task for any model to reduce feature redundancy. The bat algorithm has been widely adopted for optimizing feature selection and classifier parameters in IDS. For instance, a mutative scale chaotic bat (MSCB) algorithm with a back-propagation neural network improved detection accuracy on datasets, such as UNSW-NB15 and KDD-Cup-99 (Xu, Qin, and Zhou 2022). Similarly, a multi-objective binary bat technique optimized feature subsets for multi-layer perceptron (MLP) based IDS, leading to better performance and reduced false positives (Ghanem et al. 2022).

Optimizing Machine Learning (ML) classifiers using the bat algorithm has yielded promising results in IDS in some studies. A binary bat method with lévy flights technique was applied to fine-tune Support Vector Machine (SVM) parameters, performing 95.05% detection accuracy on the NSL-KDD dataset, with a significantly reduced false alarm rate (Enache and Sgârciu 2014).

Feature selection using the bat algorithm has enhanced detection performance in various cybersecurity applications. The bat algorithm optimizes input features for ML classifiers, achieving a 97.40% detection rate with a false-positive rate of only 0.029% on the KDD99 dataset (Narayanasami et al. 2022).

Hybrid optimization approaches combining bat with other techniques have improved feature selection and classification efficiency. The hybrid genetic bat-based technique was also utilized for text categorization, demonstrating its potential for selecting IDS features (Eligüzel, Çetinkaya, and Dereli 2022; Rauf et al. 2020; Yu et al. 2025).

LSTM and Deep Learning-Based IDS

LSTMs have been widely used in IDS for modeling sequential network traffic patterns (Yu et al. 2022; Kanna and Santhi 2024). For instance, an LSTM-based IDS was developed for Vehicular Ad Hoc Networks (VANETs) to differentiate legitimate messages from false emergency alerts, outperforming traditional ML models. Similarly, integrating AutoEncoders (AE) with LSTM improved detection accuracy by balancing dimensionality reduction and feature retention (Hnamte et al. 2023). Further improvements were seen when Principal Component Analysis (PCA) was fused with AE, enhancing performance on datasets such as UNSW-NB15 (Thakkar, Kikani, and Geddam 2024).

Combining LSTM with other deep-learning techniques has shown promising results. A Convolutional Neural Network (CNN) with LSTM fused hybrid IDS, where CNN ex-

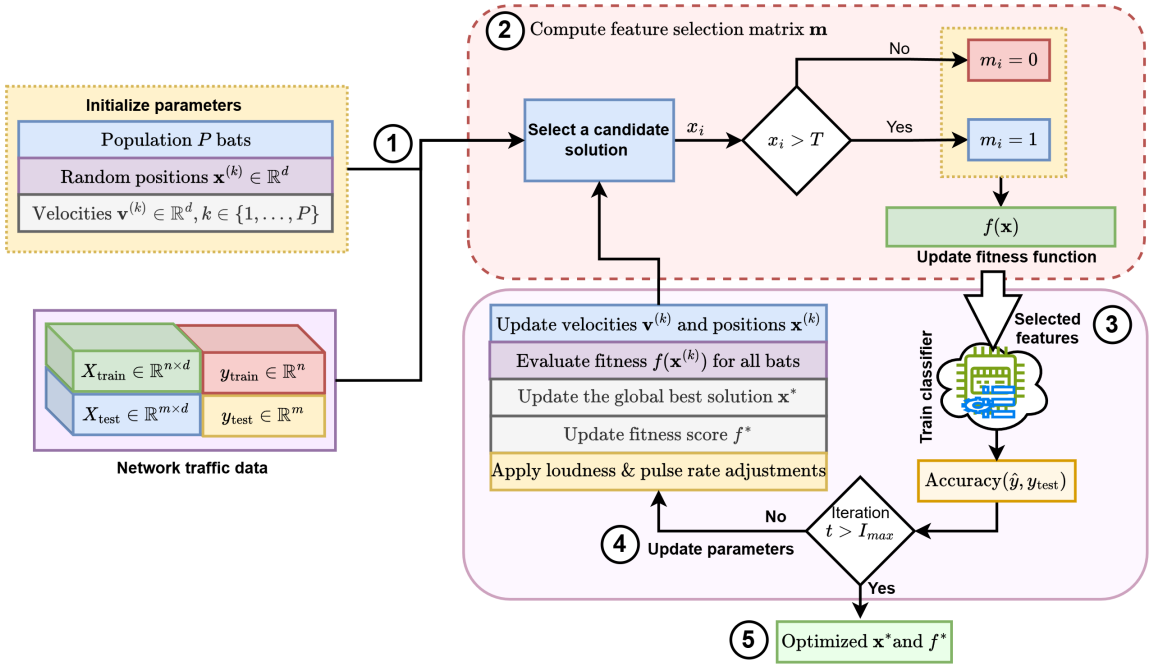

Figure 2: Illustrations of the adaptive bat optimized IDS feature selection

tracts spatial features and LSTM captures temporal dependencies, demonstrated improved detection accuracy across multiple datasets (Halbouni et al. 2022). Similarly, a BiLSTM model significantly improved the feature representation of the elaborate attack information, such as User-to-Root (U2R) and Remote-to-Local (R2L) attacks (Imrana et al. 2021).

Multi-stage IDS frameworks have been explored to enhance detection efficiency. A hierarchical IDS using AE, one-class SVM, random forests, and neural networks reduced computational overhead while improving detection accuracy (Verkerken et al. 2023; Wang et al. 2024b). While LSTM-based methods effectively model sequential patterns, they often fail to prioritize the most relevant features, leading to increased computation costs. This limitation is addressed by integrating attention mechanisms (Thakkar, Kikani, and Geddam 2024).

Attention Mechanisms in Intrusion Detection Systems

Attention mechanisms improve IDS models by narrowing the focus to the most critical features. For example, an attention-weighted spatial-temporal graph model for IoT anomaly detection demonstrated improved accuracy across multiple datasets (Wang et al. 2024a; Liu et al. 2025). Similarly, ABCNN-IDS, an attention-based CNN model, enhanced learning for low-instance attack classes, achieving 99.81% detection accuracy (Momand, Jan, and Ramzan 2024). Another attention-driven IDS utilized multi-agent models to dynamically prioritize detection agents, improving adaptability to evolving cyber threats (Sethi et al. 2021).

Integrating attention with deep-learning models has further improved intrusion detection performance. A Bidirectional Gated Recurrent Unit (Bi-GRU) and attention-fused hybrid model for Industrial Internet-of-Things (IIoT) intrusion detection effectively captured critical temporal features while mitigating class imbalance issues (Yang, Wang, and Li 2024). Similarly, combining LSTM, Recurrent Neural Network (RNN) or GRU, and attention mechanisms facilitated real-time intrusion detection (Djaidja et al. 2024), while CANET, a hierarchical CNN-attention model, optimized detection in imbalanced datasets (Ren et al. 2023).

These researches showcase the efficacy of attention-based models in refining feature optimization, reducing false positives, and improving detection accuracy. However, attention mechanisms alone are not sufficient for optimizing feature selection. This gap is addressed by integrating bat optimization to enhance feature selection and reduce computational overhead in the proposed framework, A2S-AFLNet.

LSTM-based IDS models are effective in capturing sequential network traffic patterns but often struggle with selecting the most relevant features, leading to increased computational costs (Thakkar, Kikani, and Geddam 2024). Attention mechanisms improve feature selection by prioritizing critical information (Wang et al. 2024a), while the bat algorithm optimizes feature subsets and model parameters (Xu, Qin, and Zhou 2022). However, existing studies have not fully explored the integration of these three techniques into a unified framework.

Proposed Methodology

A2S-AFLNet employs a two-stage deep-learning framework as illustrated in Figure 1. Unlike conventional LSTM models that treat all time steps equally, attention directs

Attack Class	Train	Test	Validation
Normal	47600 (32.87%)	37000 (45.87%)	8400 (32.87%)
Generic	34000 (23.48%)	18871 (23.39%)	6000 (23.48%)
Exploits	28384 (19.60%)	11132 (13.80%)	5009 (19.60%)
Fuzzers	15456 (10.67%)	6062 (7.51%)	2728 (10.67%)
DoS	10424 (7.19%)	4089 (5.07%)	1840 (7.20%)
Reconnaissance	8917 (6.16%)	3496 (4.33%)	1574 (6.16%)

Table 1: Selected class distribution for the UNSW-NB15 dataset

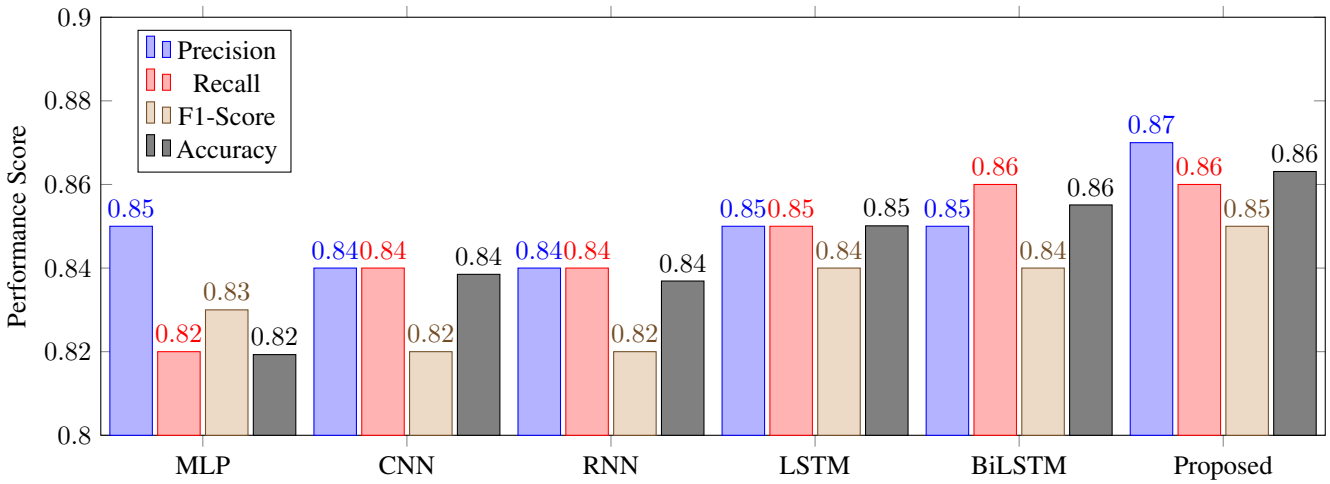

Figure 3: Performance comparisons of the proposed architecture with state-of-the-art DL models on the UNSW-NB15 dataset (No optimal feature selection)

computational focus to the most informative patterns, improving detection accuracy while reducing false positives (Hou et al. 2025). The framework uses an optimized bat algorithm to choose the best features for the IDS, which ensures only the most relevant data to inform the model while decreasing the computational overhead and improving efficiency. The IDS framework also incorporates the attention mechanism that addresses a key limitation of LSTM networks, i.e., their struggle with long-term dependencies. The attention mechanism helps in more accurate anomaly detection by highlighting important features.

Adaptive Bat for Feature Optimization

It is imperative to select an optimal feature subset as the large number of features from the network traffic data may be redundant or irrelevant for intrusion detection. To choose the optimal feature for the IDS, the hybrid bat algorithm (Lyu et al. 2019; Meng et al. 2015) is integrated with adaptive parameter tuning as illustrated in Figure 2, which enhances both global search analysis and local refinement utilization. By dynamically adjusting search criteria, this approach ensures an optimal balance between selecting the key features and minimizing computational load for further training the intrusion detection model.

Let $X_{\text{train}} \in \mathbb{R}^{n \times d}$ be the training feature matrix and $y_{\text{train}} \in \mathbb{R}^n$ be the corresponding attacks and non-attack labels. Similarly, $X_{\text{test}} \in \mathbb{R}^{m \times d}$ and $y_{\text{test}} \in \mathbb{R}^m$ represent the testing data. The feature selection process is defined as an optimization problem where a binary matrix $\mathbf{x} \in \{0,1\}^d$ determines the subset of features used for the attack classification.

A population of P bats is initialized with random positions $\mathbf{x}^{(k)} \in \mathbb{R}^d$ and velocities $\mathbf{v}^{(k)} \in \mathbb{R}^d$, where $k \in \{1, \ldots, P\}$. The feature selection matrix $\mathbf{m}$ is obtained from $\mathbf{x}$ as:

$$m_i = \begin{cases} 1, & \text{if } x_i > \text{ threshold } T, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

This step ensures diverse candidate solutions for feature selection. The selected subset $X_{\text{train},\mathbf{m}}$ is used to train a classifier $h_{\mathbf{m}}(X_{\text{train},\mathbf{m}}) \rightarrow \hat{y}_{\text{train}}$ using logistic regression. As the network traffic data contain a mix of normal and different attack classes, the logistic regression provides a better understanding of the influence for selected features using probabilistic outcome (Kolukisa et al. 2024). To ensure that the best feature subset is selected to improve intrusion detection, the fitness function is computed based on the accuracy of the classifier:

$$f(\mathbf{x}) = \alpha(1 - A) + (1 - \alpha)\frac{|\mathbf{m}|}{d}, \tag{2}$$

where A is the classification accuracy, and $|\mathbf{m}|$ represents the number of selected features.

Threshold	Feature Num	Precision	Recall	F1-Score	Accuracy (%)
0.0	42	0.85	0.85	0.84	85.04
0.02	38	0.85	0.85	0.84	85.20
0.1	**36**	**0.87**	**0.86**	**0.85**	**86.31**
0.2	31	0.85	0.85	0.84	85.15
0.3	29	0.85	0.84	0.83	84.38
0.4	25	0.85	0.85	0.84	85.27
0.5	21	0.85	0.85	0.84	84.56
0.6	15	0.80	0.80	0.79	80.11

Table 2: Optimal feature subset selection using adaptive bat on the UNSW-NB15 dataset

Attack Class	Precision	Recall	F1-Score	Accuracy (%)
Normal	0.54	0.06	0.10	92.16
Fuzzers	0.64	0.94	0.76	60.29
Reconnaissance	0.57	0.60	0.59	75.11
Exploits	1.00	0.98	0.99	94.36
DoS	0.95	0.92	0.94	5.67
Generic	0.92	0.75	0.83	97.99

Table 3: Performance of the proposed model with optimal feature selection with adaptive bat on the UNSW-NB15 dataset

The adaptive movement helps avoid local optima while ensuring the selection of an optimal feature subset from the network traffic data. The velocity is updated as follows:

$$v_i^{(k,t+1)} = v_i^{(k,t)} + \beta(x_i^* - x_i^{(k,t)}), \tag{3}$$

where x_i^* represents the global best solution and β is a random variable drawn from $[0, 1]$. The new position is given by:

$$x_i^{(k,t+1)} = x_i^{(k,t)} + v_i^{(k,t+1)}. \tag{4}$$

The bat optimized feature subset is fed into the proposed two-stage BiLSTM model with multi-head attention fusion to enhance intrusion detection.

Two-Stage Attention Fused LSTM Architecture

A two-stage BiLSTM architecture fused with attention is proposed to recognize the most relevant patterns from the network traffic for intrusion detection. The bat optimized feature subset is processed as input to the proposed architecture. This ensures that only the selected features contribute to model training, reducing dimensionality and computational overhead.

LSTMs process network traffic sequentially without attention, leading to missing critical attack indicators hidden within large datasets. The first BiLSTM with 64 units is applied to capture both forward and backward dependencies in network traffic sequences. This enhances temporal feature representation for normal and different types of attack classes. The first BiLSTM layer processes the input features to initial refined network patterns as $\mathbf{h}_1 = \text{BiLSTM}(\mathbf{X})$. The normalization is applied to compute $\mathbf{b}_1 = \text{BatchNorm}(\mathbf{h}_1)$, stabilizing the training process, accelerate convergence, and improve generalization by reducing internal covariate shifts. The multi-head attention mechanism is applied to compute $\mathbf{a}_1 = \text{MHA}(\mathbf{b}_1, \mathbf{b}_1)$, which enhances feature representation for the network traffics by assigning different attention weights to various time steps. A residual skip connection is applied to compute $\mathbf{r}_1 = \mathbf{h}_1 + \mathbf{a}_1$ preserving the initial traffic information. Finally, dropout is applied to the computed first-stage feature outcome as: $\mathbf{d}_1 = \text{Dropout}(\mathbf{r}_1)$ while reducing overfitting and improving generalization.

LSTMs process network traffic sequentially without attention, leading to missing critical attack indicators hidden within huge network traffic data. The attention-fusion allows the intrusion detection model to focus on the most contributing feature component during the training process. The attention-enhanced feature map is summed with the original LSTM output using residual connections, preserving useful information and mitigating vanishing gradient issues.

Another BiLSTM with 64 units refines the extracted feature representations and repeats the same process as the first-stage following Equation 5 to Equation 9.

$$\mathbf{h}_2 = \text{BiLSTM}(\mathbf{d}_1) \tag{5}$$

$$\mathbf{b}_2 = \text{BatchNorm}(\mathbf{h}_2) \tag{6}$$

$$\mathbf{a}_2 = \text{MHA}(\mathbf{b}_2, \mathbf{b}_2) \tag{7}$$

$$\mathbf{r}_2 = \mathbf{h}_2 + \mathbf{a}_2 \tag{8}$$

$$\mathbf{d}_2 = \text{Dropout}(\mathbf{r}_2) \tag{9}$$

The final feature representation is transformed into a dense vector suitable for classification. The output layer uses a softmax activation function to classify network traffic into multiple intrusion types, $\mathbf{y}_{\text{pred}} = \text{Softmax}(\text{Dense}(\text{Flatten}(\mathbf{d}_2)))$.

The model is trained using categorical cross-entropy loss and optimized with the Adam optimizer. The cross-entropy loss and Adam optimizer reduce the overfitting of multi-class sequential network traffic data (Ibrahim et al. 2024;

	No Attention				Attention Fused			
No. of Stage(s)	**Precision**	**Recall**	**F1-Score**	**Accuracy (%)**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
1	0.85	0.85	0.84	85.07%	0.85	0.85	0.84	84.77
2	0.86	0.85	0.85	84.95%	**0.87**	**0.86**	**0.85**	**86.31**
3	0.86	0.85	0.85	85.31%	0.81	0.81	0.80	80.90

Table 4: Performance comparisons of the proposed architecture with attention on the UNSW-NB15 dataset

Work	**Methodology**	**Size of features subset**	**F1-score (%)**	**Accuracy (%)**
(Kasongo and Sun 2020b)	ANN + XGBoost	19	77.28	77.51
(Roy and Singh 2021)	MLP + IG	20	-	84.1
(Kasongo and Sun 2020a)	FFDNN + WFEU	22	-	77.16
(Moustafa and Slay 2016)	ANN	42	-	81.34
(Yu et al. 2022)	MLP + IGRF-RFE	23	82.85	84.24
(Madwanna et al. 2023)	CNN-BiLSTM	-	-	82.19
Proposed	A2S-AFLNet	25	84	85.27

Table 5: Comparative analysis of recent studies with the proposed methodology on the UNSW-NB15 dataset

Abdelkhalek and Mashaly 2023; Kunang et al. 2021). Learning rate scheduling and early stopping mechanisms ensure efficient training and prevent further overfitting.

Results and Discussion

The proposed model, A2S-AFLNet, is evaluated using a benchmark dataset UNSW-NB15 (Moustafa and Slay 2015; Moustafa, Creech, and Slay 2017; Moustafa and Slay 2016; Moustafa, Slay, and Creech 2019; Sarhan et al. 2021) consisting normal traffic activities and synthetic contemporary attack behaviours. The performance of the A2S-AFLNet model is validated with the dataset containing six categories of attack classes, including normal behaviour of the network traffic data as described in Table 1. The normal traffic class is the largest across training (32.87%), test (45.87%), and validation (32.87%) sets. Generic and Exploits classes contribute significantly to the dataset, while Fuzzers, DoS, and Reconnaissance have smaller proportions. The A2S-AFLNet model is trained on a workstation equipped with an Intel(R) Core(TM) i7-12700 processor (2.10 GHz) and 64 GB of RAM on a Python programming environment, along with TensorFlow and Keras libraries. Also, the Colab platform (Bisong 2019) was used for partial execution and analysis of the proposed framework.

Standard performance metrics such as Accuracy, Precision, Recall, and F1-score are used to evaluate the proposed A2S-AFLNet architecture. Accuracy is computed using the proportion of correctly classified attacks and normal instances. Precision measures how many of the identified attack samples are actually correct out of all positive instances. Recall computes how the model effectively detects all actual attack samples, where fewer attack classes are expected to be misclassified as normal (false negatives). F1-score indicates that the model has a good balance between detecting all attack types while minimizing misclassification errors.

Figure 3 depicts the comparison between Deep Neural Network (DNN) models, such as MLP, CNN, RNN, LSTM, and BiLSTM, and proposed A2S-AFLNet architecture in Accuracy, Precision, Recall, F1-score. BiLSTM and LSTM achieve better Recall and Accuracy with around 0.85, while MLP, CNN, and RNN show similar performance around 0.84 for most metrics. The proposed A2S-AFLNet model achieves the highest accuracy (0.8631), Precision (0.87), Recall (0.86), and F1-Score (0.85), indicating better generalization.

Table 2 highlights the optimal feature selection and model performance. The optimal feature threshold is found to be 0.1, selecting 36 features. Feature selection improves accuracy to 86.31%, higher than other thresholds. Keeping all forty-two features (no optimal selection) results in slightly lower performance (85.04% accuracy), making some features irrelevant and affecting the model. Reducing features below 0.1 thresholds (e.g., fifteen features at 0.6 thresholds) significantly decreases accuracy to 80.11%, showing the importance of selecting an optimal subset. Table 3 shows the performance against different attacks and normal classes for the proposed architecture. The attack classes performed well in most metrics while highlighting some false positive cases in fuzzers and normal classes.

The effect of incorporating the attention mechanism is described in Table 4. Adding multi-head attention improves the performance of the proposed two-stage attention-fused model. With attention, the highest accuracy is achieved in the two-stage model (86.31%), which matches the optimal feature selection model's performance. Increases one more stage see a performance drop (80.90%) which indicates that the excessive attention stage causes overfitting or redundancy.

A comparative study of recent methodologies with our proposed framework, A2S-AFLNet, is discussed in Table 5. The A2S-AFLNet framework achieves test accuracy of 85.27% outperforming other methodologies while balancing selected feature subset size of 25 for optimal perfor-

mance. Overall, A2S-AFLNet achieved better accuracy and greater resilience against emerging threats through adaptive feature selection, attention-driven learning, and sequential modelling.

Conclusion

This study proposed A2S-AFLNet, a novel two-stage LSTM-based IDS incorporating attention to feature prioritization and the adaptive bat optimization technique to enhance feature selection and efficiency. We performed extensive experiments using the UNSW-NB15 dataset to validate A2S-AFLNet, comparing their performance with that of the leading IDS models. The results demonstrate that A2S-AFLNet surpasses traditional deep-learning and machine learning techniques, achieving higher accuracy, lower false favourable rates, and improved computational efficiency. By integrating proposed bat-selected optimal features into the two-stage LSTM-attention model, the system achieves robust and efficient intrusion detection, accurately identifying different classes of cyber threats. Future work focuses on integrating adversarial training with hybrid deep learning models optimized for resource-constrained devices to enhance robustness against more sophisticated cyber-attacks.

References

Abdelkhalek, A.; and Mashaly, M. 2023. Addressing the class imbalance problem in network intrusion detection systems using data resampling and deep learning. *The Journal of Supercomputing*, 79(10): 10611–10644.

Al-Haija, Q. A.; and Droos, A. 2025. A comprehensive survey on deep learning-based intrusion detection systems in Internet of Things (IoT). *Expert Systems*, 42(2): e13726.

Bisong, E. 2019. Google Colaboratory. In Bisong, E., ed., *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*, 59–64. Berkeley, CA: Apress. ISBN 978-1-4842-4470-8.

Djaidja, T. E. T.; Brik, B.; Mohammed Senouci, S.; Boualouache, A.; and Ghamri-Doudane, Y. 2024. Early Network Intrusion Detection Enabled by Attention Mechanisms and RNNs. *IEEE Transactions on Information Forensics and Security*, 19: 7783–7793.

Eligüzel, N.; Çetinkaya, C.; and Dereli, T. 2022. A novel approach for text categorization by applying hybrid genetic bat algorithm through feature extraction and feature selection methods. *Expert Systems with Applications*, 202: 117433.

Enache, A.-C.; and Sgârciu, V. 2014. Enhanced intrusion detection system based on bat algorithm-support vector machine. In *2014 11th International Conference on Security and Cryptography (SECRYPT)*, 1–6.

Ghanem, W. A. H. M.; Ghaleb, S. A. A.; Jantan, A.; Nasser, A. B.; Saleh, S. A. M.; Ngah, A.; Alhadi, A. C.; Arshad, H.; Saad, A.-M. H. Y.; Omolara, A. E.; El-Ebiary, Y. A. B.; and Abiodun, O. I. 2022. Cyber Intrusion Detection System Based on a Multiobjective Binary Bat Algorithm for Feature Selection and Enhanced Bat Algorithm for Parameter Optimization in Neural Networks. *IEEE Access*, 10: 76318–76339.

Halbouni, A.; Gunawan, T. S.; Habaebi, M. H.; Halbouni, M.; Kartiwi, M.; and Ahmad, R. 2022. CNN-LSTM: Hybrid Deep Neural Network for Network Intrusion Detection System. *IEEE Access*, 10: 99837–99849.

Hnamte, V.; Nhung-Nguyen, H.; Hussain, J.; and Hwa-Kim, Y. 2023. A Novel Two-Stage Deep Learning Model for Network Intrusion Detection: LSTM-AE. *IEEE Access*, 11: 37131–37148.

Hou, Y.; Wei, X.; Fan, J.; and Wang, C. 2025. Interpretable CAA classification based on incorporating feature channel attention into LSTM. *Computers & Security*, 150: 104252.

Ibrahim, S.; Youssef, A. M.; Shoman, M.; and Taha, S. 2024. Intelligent SDN to enhance security in IoT networks. *Egyptian Informatics Journal*, 28: 100564.

Imrana, Y.; Xiang, Y.; Ali, L.; and Abdul-Rauf, Z. 2021. A bidirectional LSTM deep learning approach for intrusion detection. *Expert Systems with Applications*, 185: 115524.

Kanna, P. R.; and Santhi, P. 2024. An Enhanced Hybrid Intrusion Detection Using Mapreduce-Optimized Black Widow Convolutional LSTM Neural Networks. *Wireless Personal Communications*, 138(4): 2407–2445.

Kasongo, S. M.; and Sun, Y. 2020a. A deep learning method with wrapper based feature extraction for wireless intrusion detection system. *Computers & Security*, 92: 101752.

Kasongo, S. M.; and Sun, Y. 2020b. Performance Analysis of Intrusion Detection Systems Using a Feature Selection Method on the UNSW-NB15 Dataset. *Journal of Big Data*, 7(1): 105.

Kolukisa, B.; Dedeturk, B. K.; Hacilar, H.; and Gungor, V. C. 2024. An efficient network intrusion detection approach based on logistic regression model and parallel artificial bee colony algorithm. *Computer Standards & Interfaces*, 89: 103808.

Kunang, Y. N.; Nurmaini, S.; Stiawan, D.; and Suprapto, B. Y. 2021. Attack classification of an intrusion detection system using deep learning and hyperparameter optimization. *Journal of Information Security and Applications*, 58: 102804.

Li, Y.; Li, Z.; and Li, M. 2025. A comprehensive survey on intrusion detection algorithms. *Computers and Electrical Engineering*, 121: 109863.

Liu, G.; Zhang, T.; Dai, H.; Cheng, X.; and Yang, D. 2025. ResInceptNet-SA: A Network Traffic Intrusion Detection Model Fusing Feature Selection and Balanced Datasets. *Applied Sciences*, 15(2): 956.

Lyu, S.; Li, Z.; Huang, Y.; Wang, J.; and Hu, J. 2019. Improved self-adaptive bat algorithm with step-control and mutation mechanisms. *Journal of Computational Science*, 30: 65–78.

Madwanna, Y.; B., A.; R., R. A.; and R., S. H. 2023. YARS-IDS: A Novel IDS for Multi-Class Classification. In *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)*, 1–6.

Meng, X.-B.; Gao, X. Z.; Liu, Y.; and Zhang, H. 2015. A novel bat algorithm with habitat selection and Doppler effect in echoes for optimization. *Expert Systems with Applications*, 42(17): 6350–6364.

Momand, A.; Jan, S. U.; and Ramzan, N. 2024. ABCNN-IDS: Attention-Based Convolutional Neural Network for Intrusion Detection in IoT Networks. *Wireless Personal Communications*, 136(4): 1981–2003.

Moustafa, N.; Creech, G.; and Slay, J. 2017. Big Data Analytics for Intrusion Detection System: Statistical Decision-Making Using Finite Dirichlet Mixture Models. In Palomares Carrascosa, I.; Kalutarage, H. K.; and Huang, Y., eds., *Data Analytics and Decision Support for Cybersecurity: Trends, Methodologies and Applications*, 127–156. Cham: Springer International Publishing. ISBN 978-3-319-59439-2.

Moustafa, N.; and Slay, J. 2015. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In *2015 Military Communications and Information Systems Conference (MilCIS)*, 1–6.

Moustafa, N.; and Slay, J. 2016. The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set. *Information Security Journal: A Global Perspective*, 25(1-3): 18–31.

Moustafa, N.; Slay, J.; and Creech, G. 2019. Novel Geometric Area Analysis Technique for Anomaly Detection Using Trapezoidal Area Estimation on Large-Scale Networks. *IEEE Transactions on Big Data*, 5(4): 481–494.

Narayanasami, S.; Sengan, S.; Khurram, S.; Arslan, F.; Murugaiyan, S. K.; Rajan, R.; Peroumal, V.; Dubey, A. K.; Srinivasan, S.; and Sharma, D. K. 2022. Biological Feature Selection and Classification Techniques for Intrusion Detection on BAT. *Wireless Personal Communications*, 127(2): 1763–1785.

Rauf, H. T.; Malik, S.; Shoaib, U.; Irfan, M. N.; and Lali, M. I. 2020. Adaptive inertia weight Bat algorithm with Sugeno-Function fuzzy search. *Applied Soft Computing*, 90: 106159.

Ren, K.; Yuan, S.; Zhang, C.; Shi, Y.; and Huang, Z. 2023. CANET: A hierarchical CNN-Attention model for Network Intrusion Detection. *Computer Communications*, 205: 170–181.

Roy, A.; and Singh, K. J. 2021. Multi-classification of UNSW-NB15 Dataset for Network Anomaly Detection System. In Purohit, S. D.; Singh Jat, D.; Poonia, R. C.; Kumar, S.; and Hiranwal, S., eds., *Proceedings of International Conference on Communication and Computational Technologies*, 429–451. Singapore: Springer. ISBN 9789811550775.

Sarhan, M.; Layeghy, S.; Moustafa, N.; and Portmann, M. 2021. NetFlow Datasets for Machine Learning-based Network Intrusion Detection Systems. volume 371, 117–135.

Sethi, K.; Madhav, Y. V.; Kumar, R.; and Bera, P. 2021. Attention based multi-agent intrusion detection systems using reinforcement learning. *Journal of Information Security and Applications*, 61: 102923.

Thakkar, A.; Kikani, N.; and Geddam, R. 2024. Fusion of linear and non-linear dimensionality reduction techniques for feature reduction in LSTM-based Intrusion Detection System. *Applied Soft Computing*, 154: 111378.

Verkerken, M.; D'hooge, L.; Sudyana, D.; Lin, Y.-D.; Wauters, T.; Volckaert, B.; and De Turck, F. 2023. A Novel Multi-Stage Approach for Hierarchical Intrusion Detection. *IEEE Transactions on Network and Service Management*, 20(3): 3915–3929.

Wang, X.; Wang, X.; He, M.; Zhang, M.; and Lu, Z. 2024a. Spatial-Temporal Graph Model Based on Attention Mechanism for Anomalous IoT Intrusion Detection. *IEEE Transactions on Industrial Informatics*, 20(3): 3497–3509.

Wang, Z.; Yang, X.; Zeng, Z.; He, D.; and Chan, S. 2024b. A hierarchical hybrid intrusion detection model for industrial internet of things. *Peer-to-Peer Networking and Applications*, 17(5): 3385–3407.

Xu, X.; Qin, H.; and Zhou, J. 2022. Cyber Intrusion Detection Based on a Mutative Scale Chaotic Bat Algorithm with Backpropagation Neural Network. *Security and Communication Networks*, 2022(1): 5605404.

Yang, K.; Wang, J.; and Li, M. 2024. An improved intrusion detection method for IIoT using attention mechanisms, Bi-GRU, and Inception-CNN. *Scientific Reports*, 14(1): 19339.

Yang, X.-S. 2010. Firefly Algorithm, Lévy Flights and Global Optimization. In Bramer, M.; Ellis, R.; and Petridis, M., eds., *Research and Development in Intelligent Systems XXVI*, 209–218. London: Springer. ISBN 978-1-84882-983-1.

Yu, H.; Zhang, W.; Kang, C.; and Xue, Y. 2025. A feature selection algorithm for intrusion detection system based on the enhanced heuristic optimizer. *Expert Systems with Applications*, 265: 125860.

Yu, Y.; Zeng, X.; Xue, X.; and Ma, J. 2022. LSTM-Based Intrusion Detection System for VANETs: A Time Series Classification Approach to False Message Detection. *IEEE Transactions on Intelligent Transportation Systems*, 23(12): 23906–23918.