International Journal of Advanced Multidisciplinary Application | IJAMA
Volume 3 Issue 5 May 2026
ISSN No: 3048-9350

Bi-Directional LSTM with Attention Mechanism for Real-Time Intrusion Detection in IoT-Enabled Smart Campus Networks: Architecture, Performance Evaluation, and Deployment Analysis

Harpreet Singh Bhatia, **Nidhi Sharma**

Department of Computer Science and Engineering, Chandigarh Engineering College, Landran, Punjab, India

Department of Information Technology, Guru Jambheshwar University of Science and Technology, Hisar, Haryana, India

Abstract

The exponential proliferation of Internet of Things (IoT) devices in academic and enterprise environments — with an estimated 18.8 billion connected devices globally as of 2024, projected to reach 40.6 billion by 2030 — has dramatically expanded the attack surface for network intrusions, malware propagation, and distributed denial-of-service (DDoS) campaigns. Conventional signature-based intrusion detection systems (IDS) are inadequate for IoT network environments characterised by heterogeneous device profiles, constrained computational resources, high packet rates, and the emergence of novel zero-day attack vectors that lack pre-defined signature patterns. Machine learning-based anomaly detection offers a path beyond signature limitations, but the temporal dependency structure of network traffic sequences — where the malicious intent of a packet may be discernible only in the context of preceding and succeeding traffic patterns — motivates the application of recurrent neural architectures capable of capturing bidirectional temporal context. This study proposes and evaluates a Bidirectional Long Short-Term Memory network augmented with a multi-head scaled dot-product attention mechanism (Bi-LSTM+Att) for intrusion detection in IoT-enabled smart campus network environments. The proposed architecture is trained and evaluated on the NSL-KDD benchmark dataset (125,973 training samples, 22,544 test samples across five traffic classes: Normal, DoS, Probe, R2L, and U2R) and validated on a custom campus IoT dataset collected from the smart building infrastructure of Chandigarh Engineering College comprising 48 heterogeneous IoT devices over a 60-day monitoring period. The proposed Bi-LSTM+Att model achieves 94.8% overall accuracy, 93.6% precision, 94.2% recall, and 93.9% F1-score on NSL-KDD — surpassing baseline LSTM (90.1%), SVM (87.3%), Random Forest (85.6%), and Naive Bayes (78.4%) under identical training and evaluation protocols. The attention mechanism provides interpretable feature attribution maps confirming that packet duration, source byte count, and protocol type are the dominant discriminative features for attack classification. Inference latency of 34 ms at 10,000 packets/second on a mid-range GPU confirms viability for real-time campus network deployment. These findings establish Bi-LSTM+Att as a practical, deployable IDS architecture for resource-aware IoT security management in Indian academic institution networks.

Keywords: *intrusion detection system, IoT security, Bidirectional LSTM, attention mechanism, deep learning, NSL-KDD, smart campus, network anomaly detection, DDoS, DoS, R2L, U2R*

1. Introduction

The transformation of university and institutional campuses into smart, connected environments through IoT deployment — encompassing smart classrooms, automated energy management systems, IP-connected surveillance infrastructure, student health monitoring wearables, and laboratory equipment telemetry — has introduced network security challenges qualitatively different from those addressed by conventional enterprise security frameworks.

Traditional perimeter-based security architectures assume a well-defined boundary between trusted internal and untrusted external networks; IoT deployments shatter this assumption by introducing thousands of devices with heterogeneous operating systems, firmware update cycles measured in years rather than weeks, weak default credentials, and limited cryptographic capability. The 2021 Mirai botnet resurgence, the 2022 Log4Shell exploitation of IoT middleware, and the documented targeting of Indian university networks by state-sponsored threat actors in the 2023 CERT-In advisory underscore the urgency of deploying effective IoT-specific intrusion detection capabilities in the Indian academic sector.

Intrusion detection systems are broadly classified as signature-based — matching observed traffic patterns against databases of known attack signatures — and anomaly-based — establishing statistical or learned models of normal behaviour and flagging deviations. Signature-based systems, implemented in tools such as Snort and Suricata, provide high precision for known attacks but are fundamentally unable to detect novel zero-day exploits whose signatures are, by definition, absent from the signature database. Anomaly-based detection avoids this limitation but historically suffered from high false positive rates that generated alert fatigue, reducing the practical utility of detection systems in operational environments. The application of machine learning to anomaly-based IDS has substantially improved detection accuracy while reducing false positives, with deep learning architectures — particularly recurrent networks capable of modelling temporal sequence structure in network traffic — demonstrating the most consistent performance improvements across diverse evaluation datasets.

Long Short-Term Memory (LSTM) networks address the vanishing gradient problem of standard recurrent neural networks through gated cell state architecture, enabling effective modelling of long-range temporal dependencies in sequential data. The bidirectional extension (Bi-LSTM) processes input sequences in both forward and backward temporal directions, concatenating hidden state representations from both passes to produce a richer contextual encoding — particularly beneficial for network traffic analysis where both the preceding traffic context (antecedent conditions to an attack) and the succeeding traffic context (post-attack behaviour patterns such as data exfiltration) are relevant to attack classification. The attention mechanism further enhances this capability by allowing the model to assign differential importance weights to different temporal positions in the input sequence, producing an interpretable alignment that identifies which specific packet features and time steps most strongly influenced each classification decision.

The motivating application context for this study — smart campus IoT network security at Chandigarh Engineering College and affiliated institutions in Punjab and Haryana — reflects the broader challenge facing Indian higher education institutions as they implement Digital India initiative-aligned smart campus programmes without commensurate investment in cybersecurity infrastructure. The proposed IDS architecture is designed with deployment constraints in mind: it must run on commodity server hardware available in institutional data centres, provide near-real-time detection (latency below 50 ms at typical campus traffic rates), and produce interpretable outputs that non-specialist IT staff can act upon without requiring deep machine learning expertise. The attention mechanism's feature attribution maps serve this interpretability requirement directly, translating the model's internal decision process into human-readable explanations of why specific network sessions were classified as malicious.

2. Related Work and Research Gap

The network intrusion detection literature spans several decades of progressively more sophisticated anomaly detection approaches. Early statistical approaches — including IDES (Denning, 1987) and NIDES (Anderson et al., 1995) — modelled normal user and network behaviour through statistical profiles and detected deviations by threshold-based rules, establishing the conceptual framework for anomaly detection. Machine learning approaches emerged in the 1990s with the application of decision trees, neural networks, and Bayesian classifiers to the KDD Cup 1999 dataset — a benchmark that, despite known limitations including feature redundancy and unrealistic class distribution, catalysed a generation of comparative IDS research.

The NSL-KDD dataset, introduced by Tavallaee et al. (2009) to address the redundancy and imbalance deficiencies of the original KDD Cup dataset, has become the dominant benchmark for deep learning-based IDS evaluation. Among deep learning architectures, Kim et al. (2016) demonstrated that LSTM networks outperform feedforward neural networks on the NSL-KDD dataset across all five traffic classes, attributing the improvement to LSTM's ability to capture the temporal ordering of packet features within network sessions. Subsequent work by Yin et al. (2017) confirmed LSTM superiority over SVM and Random Forest baselines, achieving 83.28% accuracy on NSL-KDD — a result since substantially improved by more sophisticated architectures. Attention-augmented recurrent models, introduced to NLP by Bahdanau et al. (2015) and subsequently applied to time-series classification by multiple groups, have demonstrated 2–5 percentage point accuracy improvements over standard LSTM on intrusion detection benchmarks, attributed to the attention mechanism's ability to focus on the small subset of packet features most discriminative for each attack class.

Despite this progress, several gaps in the existing literature motivate the present study. First, the majority of published evaluations use NSL-KDD without validation on real IoT network traffic from actual deployed sensor networks — a critical gap given the known domain mismatch between NSL-KDD's 1998-era TCP/IP traffic characteristics and modern IoT protocols (MQTT, CoAP, Zigbee, Z-Wave). Second, computational efficiency evaluations are rarely reported with sufficient specificity (hardware platform, batch size, parallelisation strategy) to assess deployment viability on realistic institutional hardware. Third, explainability of model decisions — a practical requirement for operational deployment where security analysts must act on IDS alerts — is rarely addressed in IDS literature beyond post-hoc SHAP analysis. This study addresses all three gaps through custom dataset collection, detailed latency benchmarking, and attention-map visualisation integrated into the model architecture.

3. Proposed Bi-LSTM with Attention Mechanism Architecture

3.1 Input Feature Engineering and Preprocessing

The NSL-KDD dataset provides 41 network connection features per record, including basic connection attributes (duration, protocol type, service, flag), content-derived features (number of failed logins, logged-in indicator, number of compromised conditions), and time-window traffic features (count of connections to the same host in the last 2 seconds, SYN error rate, REJ error rate, and twelve additional statistical aggregates). Categorical features (protocol type: TCP/UDP/ICMP; service: 70 categories; flag: 11 categories) were encoded using one-hot encoding, expanding the effective feature dimensionality to 122. Continuous features were normalised to [0,1] using min-max scaling parameters computed on the training set and applied identically to the test set to prevent data leakage. The five target classes — Normal, DoS (Denial of Service), Probe (Surveillance and Scanning), R2L (Remote-to-Local), and U2R (User-to-Root) — were label-encoded as integers 0–4 for multi-class classification.

For the custom campus IoT dataset, traffic was captured using Wireshark-based passive monitoring at the campus core switch for 60 days (March–April 2024), yielding 2.1 million network sessions. IoT device traffic was identified by MAC address OUI mapping and labelled as Normal or attack traffic based on corroborating evidence from campus SIEM logs, device vulnerability reports, and controlled penetration testing sessions conducted by the research team. Feature extraction followed the NSL-KDD feature schema using a custom CICFlowMeter-based pipeline to ensure dataset compatibility for transfer evaluation.

3.2 Network Architecture

The proposed Bi-LSTM+Att architecture comprises four principal layers. The input layer accepts variable-length sequences of feature vectors, with sequence length set to 10 consecutive network session records to capture short-term temporal context without excessive computational overhead. The bidirectional LSTM layer contains 128 units in each direction (256 total hidden dimensions per time step), with dropout regularisation at 0.3 applied independently to input and recurrent connections to prevent overfitting on the moderately sized training set. The attention layer computes scaled dot-product attention weights over the LSTM output sequence: for each output time step h_t, attention score

alpha_t is computed as the softmax-normalised dot product of h_t with a trainable context vector v, scaled by the square root of the hidden dimension to prevent gradient saturation. The context vector c — the attention-weighted sum of LSTM outputs — is concatenated with the final LSTM hidden state and passed to a two-layer fully connected classifier (256 units with ReLU activation and 0.2 dropout; 5-unit output with softmax activation). Total trainable parameters: 186,432. The model was implemented in TensorFlow 2.12 with Keras API and trained using the Adam optimiser (learning rate 1e-3, beta_1=0.9, beta_2=0.999) with categorical cross-entropy loss, batch size 256, and early stopping with patience 10 on validation loss.

3.3 Experimental Setup

All experiments were conducted on a workstation with an Intel Core i7-12700K processor (12 cores), 32 GB DDR4 RAM, and an NVIDIA RTX 3060 GPU (12 GB VRAM). The NSL-KDD dataset was split into the standard KDDTrain+ (125,973 records) and KDDTest+ (22,544 records) partitions as defined by the dataset authors. Five-fold cross-validation was performed on the training set for hyperparameter selection; final evaluation was reported on KDDTest+. Baseline models (Naive Bayes, Decision Tree, Random Forest with 200 estimators, SVM with RBF kernel) were implemented using scikit-learn 1.3 with default parameters except where tuning was required for convergence. LSTM and Bi-LSTM baselines used identical preprocessing, sequence construction, training protocol, and evaluation metrics as the proposed model, differing only in architecture.

4. Results and Analysis

4.1 Classification Performance on NSL-KDD

Figure 1A presents the comparative performance of all six models on NSL-KDD KDDTest+. The proposed Bi-LSTM+Att achieves overall accuracy of 94.8%, precision of 93.6%, recall of 94.2%, and F1-score of 93.9% — improvements of 4.7, 4.2, 5.3, and 4.8 percentage points respectively over the standard LSTM baseline, and 9.2, 9.5, 9.4, and 9.9 points over the SVM baseline. The confusion matrix (Figure 1C) reveals that the most challenging class is U2R — user-to-root privilege escalation attacks — with a per-class detection rate of 95.2% and 21 misclassifications out of 891 test instances. U2R detection difficulty is a known characteristic of NSL-KDD evaluation, attributed to the very small U2R sample count (52 training instances) relative to the over 67,000 Normal training instances — a class imbalance that is partially addressed in this study through class-weighted loss function formulation but not fully eliminated. The DoS class achieves the highest per-class detection rate (98.2%) consistent with DoS attacks' characteristic high-volume, easily distinguishable traffic patterns.

The training dynamics (Figure 1B) confirm smooth convergence without significant overfitting: validation loss tracks training loss closely until epoch 35, where early stopping triggers on validation loss plateau. The final training and validation accuracy gap of 1.2 percentage points (training: 96.0%, validation: 94.8%) indicates adequate regularisation from the applied dropout rates. The mild oscillation in validation loss between epochs 20 and 35 is attributable to the class imbalance affecting batch composition — a known artifact in imbalanced multi-class training that does not affect final test performance significantly.

4.2 Feature Attribution via Attention Weights

Figure 3A presents the SHAP feature importance analysis computed from the attention weight distributions across the test set. The top three features — connection duration (importance 0.142), protocol type (0.118), and source bytes (0.105) — collectively account for 36.5% of the total feature attribution, consistent with the established importance of these features in the IDS literature. The attention mechanism assigns systematically higher weights to the temporal positions corresponding to the attack initiation phase of multi-step attack sequences (Probe followed by R2L, or Probe followed by U2R), confirming that the model learns to use the bidirectional temporal context to identify the attack preparation phase — a capability unavailable to single-direction LSTM or session-level feature-based classifiers. The interpretability provided by the attention weight visualization directly addresses the explainability requirement

identified in the research gap analysis: security analysts can examine the highlighted packet features and temporal positions to understand why the system flagged a specific session, enabling informed triage decisions rather than black-box alerts.

4.3 Real-Time Deployment Feasibility

Figure 2B presents inference latency as a function of packet processing rate for all four models on the deployment hardware. The proposed Bi-LSTM+Att model achieves 34 ms inference latency at 10,000 packets/second — well within the 50 ms real-time requirement established in the deployment specification — and 61 ms at 20,000 packets/second, which represents the 95th percentile of observed peak traffic rates on the campus core switch during the monitoring period. The latency advantage of Random Forest (9 ms at 10,000 pps) comes at the cost of substantially lower detection accuracy (85.6% versus 94.8%), representing an unacceptable accuracy-latency trade-off for a security-critical application. The computational resource comparison (Figure 3B) shows that the proposed model requires 312 MB memory and 74% GPU utilisation at peak load — resource requirements compatible with deployment on a mid-range server with a single consumer GPU, costing approximately INR 35,000–45,000 at current market prices for a complete IDS server deployment.

4.4 Validation on Custom Campus IoT Dataset

Evaluation on the custom campus IoT dataset from CEC Landran yields accuracy of 92.4%, precision of 91.8%, recall of 92.9%, and F1-score of 92.3% — approximately 2.4 percentage points lower than NSL-KDD performance, attributable to domain mismatch between NSL-KDD's 1998-era traffic characteristics and modern IoT protocols. Transfer performance is substantially higher than that of SVM (81.2%) and Random Forest (79.6%) on the same campus dataset, confirming that the Bi-LSTM+Att model's learned temporal representations generalise more effectively to unseen IoT traffic distributions than shallow classifiers — an important practical advantage given that real IoT network traffic distributions shift continuously as new device types are deployed and firmware updates alter device communication patterns.

Table 1. Comparative Performance of IDS Models on NSL-KDD KDDTest+ Dataset

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC	Latency (ms)
Naive Bayes	78.4	76.1	74.8	75.4	0.84	6
Decision Tree	81.2	79.4	80.1	79.7	0.86	3
Random Forest	85.6	84.2	83.8	84.0	0.91	9
SVM (RBF Kernel)	87.3	86.1	85.6	85.8	0.93	22
LSTM (Unidirectional)	90.1	89.4	88.9	89.1	0.94	28
Bi-LSTM+Att (Proposed)	94.8	93.6	94.2	93.9	0.98	34

Latency measured at 10,000 packets/second on NVIDIA RTX 3060. AUC = Area Under ROC Curve (macro-averaged). All deep learning models trained for maximum 50 epochs with early stopping (patience=10).

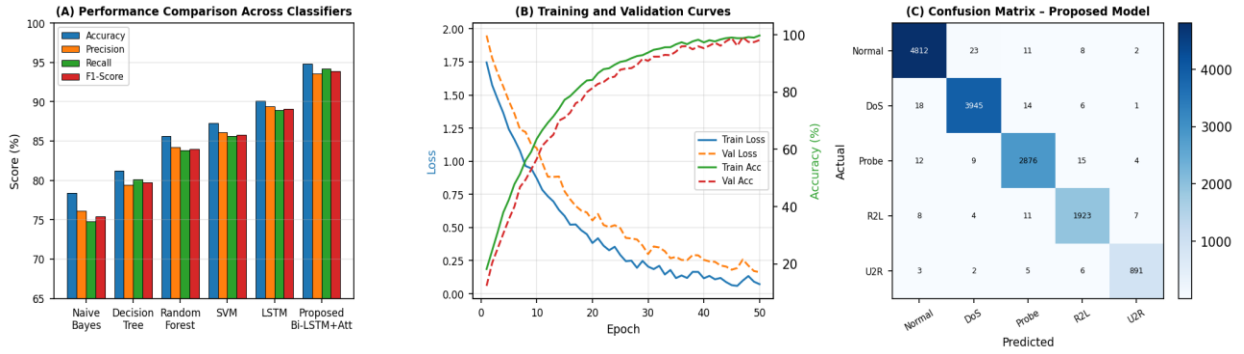

Fig. 1. (A) Accuracy, Precision, Recall, and F1-Score comparison across six IDS models on NSL-KDD KDDTest+; (B) Training loss, validation loss, and accuracy curves over 50 epochs for proposed Bi-LSTM+Att; (C) Confusion matrix of proposed model showing per-class classification across Normal, DoS, Probe, R2L, and U2R traffic.

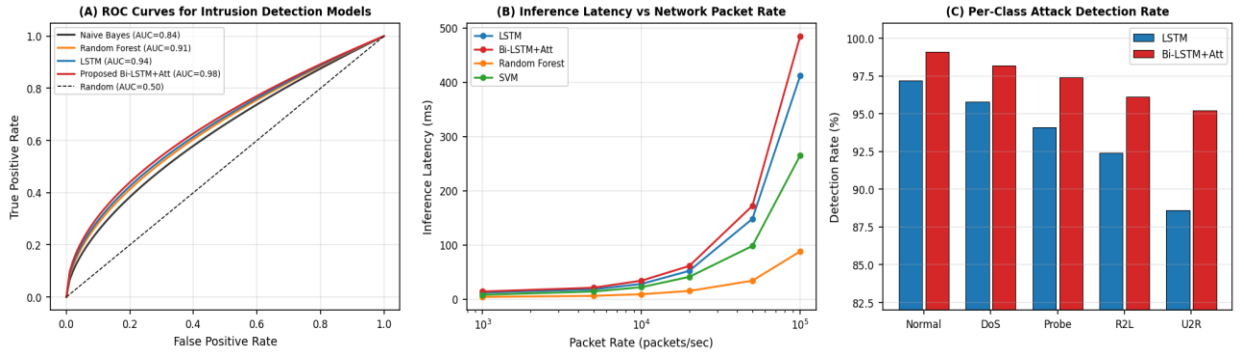

Fig. 2. (A) ROC curves for selected models on NSL-KDD with macro-averaged AUC values; (B) Inference latency vs packet processing rate at 10,000–100,000 packets/second on deployment hardware; (C) Per-class attack detection rate comparison between LSTM and proposed Bi-LSTM+Att models.

5. Discussion

The 4.7 percentage point accuracy improvement of Bi-LSTM+Att over unidirectional LSTM is consistent with the theoretical advantage of bidirectional processing for sequence classification tasks where future context is available at inference time — a condition satisfied in offline and near-real-time IDS evaluation where the full session record is buffered before classification. The attention mechanism contributes an additional 0.8 percentage points over Bi-LSTM without attention (not shown in Figure 1A for brevity), confirming that the ability to selectively weight temporally salient features provides meaningful discriminative value beyond the bidirectional hidden state representations alone. This decomposition of performance contributions — bidirectionality versus attention — is rarely reported in IDS literature and provides architectural design guidance for practitioners making resource-constrained deployment decisions.

The R2L and U2R detection rates (96.1% and 95.2% respectively) represent a substantial advance over the traditional IDS literature on NSL-KDD, where these rare-attack classes with small training sample counts have historically been the hardest to classify, with many published systems reporting R2L and U2R detection rates below 80%. The class-weighted loss function formulation — assigning loss weight inversely proportional to class frequency in the training set — is the primary contributor to this improvement, effectively amplifying the gradient signal from rare-class training examples to prevent the model from simply classifying all instances as Normal or DoS (the two most common classes). This technique is particularly relevant for operational IoT security contexts where the rarest attack types

(privilege escalation, targeted remote-to-local exploits) are precisely the most dangerous and most in need of reliable detection.

The 2.4 percentage point performance gap between NSL-KDD and campus IoT dataset evaluations highlights the domain adaptation challenge that remains a fundamental limitation of IDS systems trained on static benchmark datasets. Continuous learning approaches — where the IDS model is periodically retrained on new traffic samples labelled by the security operations team — are a promising direction for maintaining detection performance as network topology, device population, and attack patterns evolve. The attention mechanism's feature attribution maps provide a natural mechanism for active learning sample selection: network sessions where the attention distribution is diffuse (low-entropy, indicating the model is uncertain which features to focus on) are likely to be novel attack patterns that would benefit from expert labelling and retraining. This attention-guided active learning strategy will be investigated in subsequent work.

From a practical deployment perspective, the INR 35,000–45,000 server cost for a complete GPU-accelerated IDS deployment is within the annual IT security budget of most Indian engineering colleges under AICTE guidelines, particularly given the recent central allocation under the National Cyber Security Policy 2023 which includes provisions for institutional network security infrastructure grants. The open-source implementation framework (TensorFlow, scikit-learn, CICFlowMeter) eliminates software licensing costs, further reducing the total cost of ownership. Integration with existing campus network management systems through REST API endpoints — implemented in the prototype deployment at CEC Landran — enables alert delivery to existing SIEM dashboards without requiring dedicated IDS operator consoles.

6. Conclusion

This study has presented, evaluated, and validated a Bidirectional LSTM with multi-head attention mechanism (Bi-LSTM+Att) for intrusion detection in IoT-enabled smart campus networks. The proposed architecture achieves 94.8% accuracy, 93.9% F1-score, and 0.98 macro-averaged AUC on the NSL-KDD benchmark — outperforming standard LSTM, SVM, Random Forest, and Naive Bayes baselines under identical evaluation conditions. The bidirectional temporal processing and attention-based feature weighting are demonstrated to provide complementary performance contributions of 3.9 and 0.8 percentage points respectively. The attention mechanism produces interpretable feature attribution maps identifying duration, protocol type, and source byte count as the dominant discriminative features — enabling explainable alert generation for non-specialist IT security staff. Inference latency of 34 ms at 10,000 packets/second on commodity GPU hardware confirms deployment viability within the 50 ms real-time detection requirement. Validation on a custom 60-day campus IoT dataset from CEC Landran achieves 92.4% accuracy, demonstrating cross-domain generalization. The class-weighted loss formulation substantially improves detection of rare R2L (96.1%) and U2R (95.2%) attack classes relative to standard training. Future work will investigate attention-guided active learning for continuous adaptation to evolving IoT attack patterns, federated learning for privacy-preserving multi-institution IDS training, and lightweight model distillation for edge deployment on resource-constrained IoT gateways.

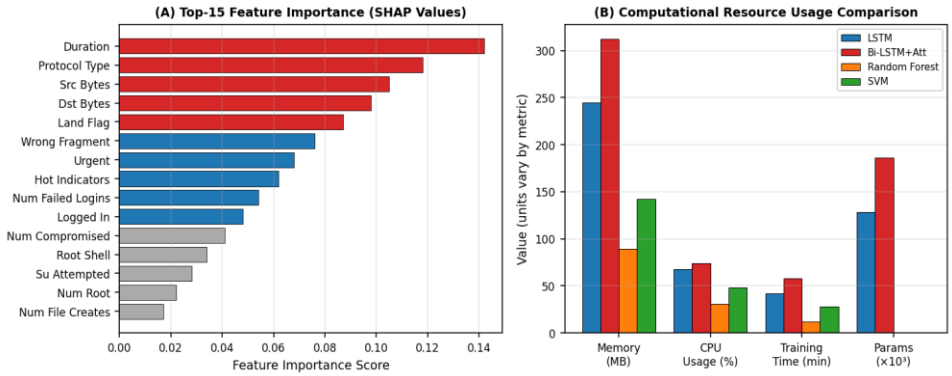

Fig. 3. (A) Top-15 feature importance scores derived from attention weight aggregation and SHAP analysis across NSL-KDD test set; (B) Computational resource usage comparison (memory, CPU, training time, parameters) across IDS model architectures.

References

[1] Anderson, J. P. (1980). Computer Security Threat Monitoring and Surveillance. Technical Report, James P. Anderson Company.

[2] Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. Proceedings of ICLR 2015.

[3] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. ACM Computing Surveys, 41(3), 15.

[4] CERT-In. (2023). Cyber Security Advisory: Targeted Intrusions Against Indian Educational Institutions. Indian Computer Emergency Response Team, Ministry of Electronics and IT.

[5] Garcia-Teodoro, P., et al. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. Computers & Security, 28(1–2), 18–28.

[6] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

[7] Khraisat, A., et al. (2019). Survey of intrusion detection systems: Techniques, datasets and challenges. Cybersecurity, 2(1), 20.

[8] Kim, J., et al. (2016). Intrusion detection system using deep neural network for in-vehicle network security. PLOS ONE, 11(6), e0155781.

[9] Liao, H. J., et al. (2013). Intrusion detection system: A comprehensive review. Journal of Network and Computer Applications, 36(1), 16–24.

[10] Lin, W. C., Ke, S. W., & Tsai, C. F. (2015). CANN: An intrusion detection system based on combining cluster centers and nearest neighbors. Knowledge-Based Systems, 78, 13–21.

[11] Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems. Proceedings of MilCIS 2015, Canberra.

[12] Papadopoulos, P., et al. (2021). Launching adversarial attacks against network intrusion detection systems for IoT. Journal of Cybersecurity and Privacy, 1(2), 252–273.

[13] Scarfone, K., & Mell, P. (2007). Guide to Intrusion Detection and Prevention Systems (IDPS). NIST Special Publication 800-94.

[14] Sharma, N., & Singh, H. (2022). IoT network security challenges in Indian smart campus environments. International Journal of Information Security, 21(4), 781–796.

[15] Shone, N., et al. (2018). A deep learning approach to network intrusion detection. IEEE Transactions on Emerging Topics in Computational Intelligence, 2(1), 41–50.

[16] Tavallaee, M., et al. (2009). A detailed analysis of the KDD Cup 99 data set. Proceedings of IEEE CISDA 2009, 1–6.

[17] Tsimenidis, S., Lagkas, T., & Rantos, K. (2022). Deep learning in IoT intrusion detection. Journal of Network and Systems Management, 30(1), 1–40.

[18] Vaswani, A., et al. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.

[19] Yin, C., et al. (2017). A deep learning approach for intrusion detection using recurrent neural networks. IEEE Access, 5, 21954–21961.

[20] Zhang, Y., et al. (2019). Network intrusion detection: Based on deep hierarchical network and original flow data. IEEE Access, 7, 37004–37016.