Anusandhanvallari
Vol 2024, No.1
March 2024
ISSN 2229-3388

Sequence-Based Threat Identification in Modern Networks Via Recurrent Memory Architectures and Game-Theoretic Feature Verification

Aravind Chagantipati

Department of Computer Science and Engineering, Kennedy University

Abstract: Contemporary network infrastructure faces sophisticated, multi-phase attack campaigns such as stealthy reconnaissance and Advanced Persistent Threats (APTs) that unfold deliberately across elongated time frames. Traditional point-in-time machine learning approaches parse network transactions as discrete, isolated entities, neglecting the vital directional dependencies embedded within traffic sequences. This investigation introduces a robust sequence-aware intrusion detection framework leveraging recurrent Long Short-Term Memory (LSTM) neural networks. When evaluated against the benchmark UNSW-NB15 dataset, the architecture dynamically captures deep historical dependencies hidden across contiguous communication sequences. Empirical execution demonstrates optimal performance, yielding a 97.5% classification accuracy and an F1-score of 97.0%. To guarantee model transparency and alignment with security operations, a post-hoc diagnostic layer utilizing game-theoretic attributions translates the complex parameter gates of the network into explicit, human-interpretable feature importance metrics at the connection layer.

Keywords: *Long Short-Term Memory, Sequence Analysis, Network Anomalies, Transparent AI, SHAP Framework, Threat Remediation.*

1. INTRODUCTION

Modern enterprise networks operate within a hostile landscape marked by multi-layered, highly integrated malicious behaviors. In contrast to legacy volumetric attacks that trigger immediate, noticeable traffic shifts, contemporary threat paradigms—including stealth scans and Advanced Persistent Threats (APTs)—manifest with extreme discretion. These sophisticated attack patterns deliberately fragment their hostile footprints across protracted temporal windows, placing innocent traffic intervals between successive anomalies. If network traffic is evaluated strictly as independent records by conventional static machine learning systems, these isolated events appear entirely normal, allowing adversaries to penetrate critical perimeters without triggering security alerts.

To overcome this detection gap, modern intrusion monitoring platforms must shift toward sequential pattern modeling, interpreting continuous network activities as inter-dependent time-series streams. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) models, are mathematically optimized to track complex temporal sequences across sequential datasets. Utilizing specialized memory cells governed by internal gating units, LSTMs successfully retain contextual historical data over extended traffic sequences, making them highly effective at identifying the subtle signatures of multi-stage attacks. Nonetheless, the dense structural parameters of recurrent models introduce significant black-box opacity, hindering real-world operational trust. This study addresses both challenges by detailing an accurate, interpretable sequence-aware network monitoring pipeline that delivers robust detection rates on the UNSW-NB15 benchmark while generating clear feature attribution metrics using cooperative game theory.

2. LITERATURE REVIEW & HISTORICAL FOUNDATIONS

Prior research in temporal data modeling establishes that sequence-aware architectures outperform traditional static methods when analyzing time-varying data streams. Within the network security domain, initial implementations utilized standard RNN loops to inspect continuous system logs, demonstrating notable

improvements over basic models like traditional Decision Trees. However, standard RNN architectures struggle with the vanishing gradient problem, which restricts their effective tracking capability to very short windows. Integrating specialized gated structures within LSTMs resolves this limitation, allowing deep tracking over extensive series of packet streams. Despite these structural advances, modern security literature frequently fails to provide rigorous validation layers that map the interior parameters of recurrent models into intuitive, actionable feature importance metrics for security practitioners.

3. RECURRENT MEMORY NETWORK FORMULATION

The analytical core of the proposed framework relies on an LSTM block designed to sequentially process multi-dimensional feature sequences x_t while dynamically maintaining historical state histories C_t across discrete processing intervals. The internal mathematical transitions occurring within the recurrent node at sequence step t

are formulated as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)\ i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)\ h_t = o_t \cdot tanh(C_t)$$

In this system, f_t, i_t, and o_t designate the output vectors produced by the forget, input, and output parameter gates, respectively. The matrices denoted by W and vectors denoted by b indicate the trainable weights and operational biases corresponding to each layer. The notation σ defines the non-linear sigmoid activation mechanism, C_t acts as

the revised internal cell state tracking block, and h_t specifies the final hidden state vector transmitted concurrently

to the subsequent temporal node and the terminal classification layer.

4. METHODOLOGY AND DATA PREPROCESSING

The sequence-based network monitoring model was trained and validated using the diverse UNSW-NB15 dataset, which covers various modern attack classes including exploits, fuzzers, backdoors, and reconnaissance maneuvers. The data engineering pipeline restructured the tabular connection streams into contiguous, overlapping windows with a uniform length of $T = 10$ sequential records. Continuous numeric variables were scaled using min-max normalization to map all entries into a uniform range of [0, 1], whereas categorical elements were converted into sparse one-hot encoded representations. Network optimization was conducted across 40 full training epochs using the Adam optimization algorithm to minimize binary cross-entropy loss, incorporating a dropout coefficient of 0.2 between layers to control over-fitting.

5. EVALUATION AND PERFORMANCE ANALYSIS

Following architecture optimization, the recurrent framework was tested on the unseen test partition of the UNSW-NB15 dataset. Table 1 lists the specific quantitative performance metrics achieved by the sequence-aware framework.

Table 1: Quantitative Evaluation Metrics for the Sequence-Aware LSTM Model

Model Type	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Inference Overhead
Long Short-Term Memory (LSTM)	97.5	96.9	97.2	97.0	3.4

ms / sample

The empirical results provided in Table 1 demonstrate that the recurrent model provides robust classification capabilities, achieving a high accuracy rate of 97.5% alongside a balanced F1-score of 97.0%. These outcomes support the hypothesis that capturing temporal context across network connection histories yields far better classification capability than relying on static, point-in-time machine learning approaches. This sequence-tracking capability allows the architecture to accurately identify complex, multi-stage threats that exhibit only minor statistical variations over extended periods.

6. POST-HOC INTERPRETABILITY AND LOCAL ATTRIBUTIONS

To mitigate the structural black-box limitations inherent in deep recurrent networks, a post-hoc Kernel SHAP diagnostic engine was integrated directly into the inference pipeline, allowing the calculation of explicit feature attributions for sequential records. Global attribution assessments identified that ***ct_srv_src*** (the frequency of connections sharing identical source locations and target services) along with ***smean*** (the average payload size sent by the source) exert the highest influence on the classification output. Individual instance-level evaluations for an identified reconnaissance alert explicitly highlighted how malicious indicators accumulated across the sequential observation window. The visualization confirmed that an incremental expansion in connection frequency across consecutive steps pushed the network's predictive score past the anomaly cutoff, offering security personnel a clear, verifiable rationale for the generated alarm.

7. CONCLUSION

This research demonstrates that sequence-aware modeling using Long Short-Term Memory networks provides a highly accurate, robust framework for detecting multi-stage network threat signatures. By coupling this architecture with post-hoc game-theoretic SHAP attributions, we show that the traditional interpretability challenges of deep recurrent networks can be effectively resolved. This dual approach delivers both high-performance threat detection and clear, auditable feature attributions, allowing security teams to confidently deploy advanced deep learning models within real-world production environments.

REFERENCES

[1] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.

[2] M. Ring et al., "A survey of network-based intrusion detection data sets," *Computers & Security*, vol. 86, pp. 147-167, 2019.

[3] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.