ISSN: 2424-8460(O)
2251-2608(P)

Original Research Article

Research on intelligent network traffic classification and dynamic scheduling mechanism based on deep learning

Ming Tang, Wenting Ding, Hongxiu Duan
Communication University of China, Nanjing, Nanjing, Jiangsu, 211172, China

Abstract: The explosive growth of network traffic and increasing complexity of service types have rendered traditional classification and scheduling methods inadequate for meeting efficient transmission demands. By leveraging convolutional neural networks and long short-term memory networks as core architectures, combined with attention mechanisms, this study proposes an intelligent traffic classification model that automatically extracts spatiotemporal traffic features, enabling precise identification of encrypted and complex traffic patterns. A dynamic scheduling mechanism is designed based on classification results to optimize resource allocation and transmission efficiency. Experimental results demonstrate that this approach significantly improves classification accuracy and network throughput, providing robust technical support for intelligent network management and security protection.
Keywords: deep learning; network traffic classification; attention mechanism; dynamic scheduling; feature extraction

1. Introduction

The widespread adoption of digital technologies has characterized network traffic with high concurrency, diverse types, and robust encryption. Traditional classification methods based on ports and manual features exhibit low accuracy and poor adaptability, while static scheduling strategies struggle to meet real-time service demands. Deep learning, leveraging its strengths in autonomous feature learning and pattern recognition, offers a novel approach for intelligent traffic processing. Deep learning models integrated with attention mechanisms can efficiently capture critical traffic features, enabling precise classification and dynamic scheduling decisions that optimize network resource utilization and ensure service quality—Making this approach the core direction for next-generation intelligent network management.

2. Analysis of the limitations of traditional network traffic classification and scheduling mechanisms

Traditional network traffic classification primarily relies on port matching, deep packet inspection, and manual feature extraction, which exhibit significant limitations in high-speed heterogeneous networks. Port matching struggles to handle dynamic, random ports, and encrypted traffic, failing to ensure reliable identification accuracy; deep packet inspection depends on fixed feature libraries, cannot analyze encrypted or unknown protocol traffic, and its computational complexity increases dramatically with rising bandwidth, making it inadequate for line-rate processing; traditional machine learning-based classification relies on manually designed features, resulting in poor generalization capabilities and a substantial decline in accuracy during traffic distribution changes or service spikes [1]. Current scheduling mechanisms predominantly employ static queues and fixed weight allocation models, lacking adaptive regulation based on traffic types, latency requirements, and link conditions. This leads to high-priority traffic bandwidth being preempted while low-priority resources are wasted, and frequently causes queue congestion and transmission jitter. Furthermore, traffic classification and scheduling modules operate in isolation without closed-loop coordination, preventing real-time optimization of scheduling strategies. Consequently, the system responds slowly and fails to adapt to high-traffic, low-latency, multi-service concurrent network scenarios, thereby hindering high-performance transmission and efficient resource utilization.

ISSN: 2424-8460(O)
2251-2608(P)

3. Construction of an intelligent traffic classification model based on deep learning

3.1. Flow data preprocessing and feature normalization

The original network traffic contains multi-dimensional heterogeneous data including packet length, arrival intervals, protocol types, and session timing, which requires standardization to accommodate deep learning model inputs. For continuous numerical features, maximum-minimum normalization is applied to map the data to the [0,1] interval, eliminating dimensionality differences. For temporal traffic patterns, fixed-length sliding windows are used to segment sessions, ensuring consistent input sequence length [2]. Discrete protocol features are converted into high-dimensional sparse vectors via one-hot encoding and concatenated with temporal features to form the model input tensor. The preprocessing workflow removes malformed packets, duplicate packets, and empty sessions, reducing noise interference on feature learning and enhancing model convergence speed and classification stability.

3.2. Design of the mixed depth feature extraction network

The intelligent traffic classification model employs a fusion architecture combining CNN and LSTM with an attention mechanism, enabling joint extraction of spatial local features and long-term temporal dependency features. The CNN module consists of stacked multi-layer one-dimensional convolutional layers, batch normalization layers, and maximum pooling layers, utilizing multi-scale convolutional kernels to capture local salient features such as packet size distribution and packet interval distribution, achieving progressive abstraction from shallow to deep features[3]. The LSTM module models the dynamic temporal changes in traffic by employing gate mechanisms (forget gate, input gate, output gate) to control the retention and updating of historical information, mitigating the gradient vanishing problem in long sequences and enhancing feature representation capabilities for both continuous and bursty traffic patterns. The gating computation adheres to the following rules:

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

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

Where $f_t i_t \sigma h_{t-1} x_t$ is the output of the forgetting gate, is the output of the input gate, is the Sigmoid activation function, is the hidden state from the previous time step, and is the input at the current time step.

3.3. Attention weighting and classification output layer design

A channel attention module is introduced at the mixed feature output stage to assign adaptive weights to features across different dimensions, enhancing the contribution of key features to classification while suppressing interference from redundant features. The attention weights are computed through global pooling and two fully connected layers, then multiplied channel-wise with the original feature maps to achieve weighted enhancement. The weighted features are fed into a fully connected layer for dimensionality mapping, ultimately producing the probability distribution of various traffic categories via the Softmax function, thereby enabling multi-classification decision-making. The classification confidence is calculated as follows:

$$\hat{y}_k=\frac{exp(z_k)}{\sum_{i=1}^{K} exp\left(z_i\right)}$$

The formula $z_k k\hat{y}_k$ represents the fully connected output of the specified category, with the predicted probability corresponding to that category. This architecture enables end-to-end feature learning and classification decision-making automatically, eliminating the need for manual rule design or feature engineering. It maintains high recognition accuracy and strong generalization capability even in scenarios involving encrypted traffic or unknown service traffic[4].

4. Dynamic scheduling mechanism driven by flow classification and performance verification

4.1. Design of the dynamic scheduling mechanism

Based on deep learning traffic classification results, a multi-priority dynamic scheduling mechanism is

ISSN: 2424-8460(O)
2251-2608(P)

established to categorize services into high-, medium-, and low-priority queues. High-priority services are allocated 45% of bandwidth resources, medium-priority services 35%, and low-priority services 20%. The scheduler dynamically adjusts queue weights and scheduling intervals according to real-time classification labels and link load conditions. When link utilization exceeds 85%, congestion avoidance is activated: High-priority services receive priority forwarding and bandwidth guarantees, while low-priority services undergo rate limiting and buffer scheduling, enabling refined bandwidth resource allocation[5].

4.2. Performance verification and result analysis

The experiment utilized standard public datasets and real network traffic, with a test bandwidth of 1000 Mbps and a sample size of 50,000 packets. Evaluation metrics included accuracy, average latency, and bandwidth utilization rate. Compared with static scheduling and traditional weighted round-robin scheduling, the proposed mechanism achieved a classification accuracy of 95.7%, reduced average service latency by 28.3%, and improved bandwidth utilization to 92.1%, outperforming all comparative approaches across all metrics.

Table 1. Performance comparison of different scheduling mechanisms.

Scheduling Plan	Classification accuracy rate	Average delay /ms	Bandwidth utilization rate
static scheduling	82.4%	42.6	73.5%
Weighted Circular Scheduling	86.1%	35.1	81.8%
Dynamic Scheduling in This Article	95.7%	30.6	92.1%

The test results demonstrate that the dynamic scheduling mechanism based on deep learning classification can rapidly respond to network state changes, enhance resource utilization efficiency and service transmission quality, and meet real-time scheduling requirements in high-speed network environments.

5. Epilogue

In complex and dynamic network environments, traditional traffic classification and static resource scheduling can no longer meet the demands for efficient management. An intelligent classification model centered on deep learning and incorporating attention mechanisms automatically extracts traffic spatiotemporal features, enabling precise identification of various traffic types and encrypted traffic. A dynamic scheduling mechanism designed based on classification results allocates network resources on demand, enhancing transmission efficiency and service quality. Experimental results demonstrate significant advantages of this approach in classification accuracy and resource utilization, providing a viable technical pathway for intelligent network operations and security protection. Future efforts should focus on further optimizing model lightweight design and real-time performance to improve adaptability in large-scale heterogeneous networks.

References

[1] Zhang X, Yin J. Optimized extreme learning machines with deep learning for high-performance network traffic classification[J]. Scientific Reports, 2025, 15(1): 33199.

[2] Agurto N D, Fuertes W, Marrone L, et al. A Novel Traffic Classification Approach by Employing Deep Learning on Software-Defined Networking[J]. Future Internet, 2024, 16(5): 153-.

[3] Gaba S, Budhiraja I, Kumar V, et al. Advancements in enhancing cyber-physical system security: Practical deep learning solutions for network traffic classification and integration with security technologies.[J]. Mathematical biosciences and engineering: MBE, 2024, 21(1): 1527-1553.

[4] Liu Youbang, Deng Yongzan. Computer Network Traffic Classification and Analysis Based on Deep Learning [J]. Information Recording Materials, 2025, 26(08): 28-30.

[5] He Erlu, Wu Xiangbo, Liu Lizhe, et al. Deep learning-based network baseline and encrypted traffic classification technology [J]. Journal of Cyberspace Security Science, 2024, 2(02): 76-85.