

CHALLENGES OF AI-BASED TRAFFIC CLASSIFICATION IN MODERN TELECOMMUNICATION NETWORKS

Feruza Tojjeva

Tashkent University of Information Technologies

Abstract: The increasing use of encrypted protocols and dynamic applications in modern telecommunication networks has significantly reduced the effectiveness of traditional traffic classification methods. Although artificial intelligence and machine learning algorithms can achieve high classification accuracy using flow-based features, accuracy alone cannot fully guarantee reliable performance under real-world network conditions. This study analyzes the key challenges associated with the practical deployment of AI-based traffic classifiers. In particular, it considers dynamic changes in network traffic, data imbalance, computational complexity, and the limited interpretability of model decisions. The analysis shows that traffic classifiers for real telecommunication networks should be evaluated using a multi-criteria approach that considers not only classification accuracy but also model generalization, resource efficiency, and adaptability.

Keywords: network traffic classification, artificial intelligence, machine learning, encrypted traffic, generalization, adaptive learning, explainable artificial intelligence

In modern telecommunication networks, the volume and diversity of traffic generated by video streaming, IP telephony, messaging, cloud services, and file-sharing applications continue to increase. The growing functionality of network applications further complicates traffic identification, as a single application may generate multiple flows with different statistical and temporal characteristics [1]. For example, modern messaging applications may simultaneously support text messaging, voice and video calls, and media transfer, each producing distinct traffic patterns. Consequently, conventional traffic identification based on static port numbers or predefined application information is becoming less reliable in modern network environments.

Table 1.

Key challenges and promising approaches in ai-based network traffic classification

Key Challenge	Cause and Impact	Evaluation Criteria	Promising Approaches
Cocept Drif	Changes in traffic characteristics and application behavior over time	Generalization ability	Adaptive & Continual Learning
Unknown traffic clasess	Emergence of new applications, services, and previously unseen traffic flows	Detection & Misclassification rate	Open-Set Recognition, Incremental Learning
Computational complexity	High traffic volumes and limited computational resources	Inference latency & Throughput	Feature Selection, Lightweight Models
Model interpretability	Limited transparency of decision-making in complex ML/DL models	Explanation consistency & Trust	Explainable AI

Accurate network traffic classification is essential for Quality of Service management, resource allocation, network monitoring, and cybersecurity. Traditional approaches have relied mainly on port-based methods and Deep Packet Inspection. However, dynamic ports and encrypted protocols such as TLS 1.3 and QUIC have reduced the effectiveness of these approaches, increasing the use of machine learning models based on statistical and temporal traffic features. Despite their high classification performance, models trained on laboratory or static datasets may experience performance degradation when applied to different network environments or evolving traffic. This problem, associated with concept drift and distribution shift, highlights the importance of model generalization in practical traffic classification [2].

Dataset quality is another important factor affecting the reliability of traffic classification models. Training an effective machine learning classifier requires accurately labeled traffic samples that adequately represent different applications, services, and network conditions. In real network environments, however, traffic generated by a target application is often mixed with background traffic produced by operating system services, synchronization processes, software updates, and third-party services. This makes accurate flow labeling particularly challenging. Consequently, a model trained on mislabeled or insufficiently representative data may achieve high performance on a test dataset but fail to maintain comparable performance when deployed in operational networks.

Another major challenge is the continuous emergence of previously unseen traffic classes. Conventional supervised classifiers generally operate under a closed-set assumption, in which the classes encountered during deployment are expected to be the same as those represented in the training data. This assumption is difficult to maintain in real telecommunication networks, where new applications, services, application versions, and protocol implementations continuously emerge. Under such conditions, a closed-set classifier may incorrectly assign an unknown traffic flow to one of the known classes, resulting in misclassification.

Retraining a model from scratch whenever a new traffic class emerges can be computationally expensive and time-consuming. Therefore, incremental and continual learning approaches, which enable classifiers to incorporate new classes while retaining previously acquired knowledge, are becoming increasingly important for evolving network environments. In parallel, open-set recognition provides mechanisms for identifying traffic that does not belong to any of the known classes rather than forcing it into a predefined category [3].

Computational complexity is another important challenge in practical traffic classification. Complex deep learning models often require substantial computational and memory resources, particularly for real-time processing. In high-speed networks, classifiers must operate with low latency without becoming a processing bottleneck. Therefore, evaluation should consider not only accuracy and F1-score, but also inference latency, processing throughput, memory consumption, and feature extraction cost. In addition, complex AI and deep learning models often have limited interpretability, making it difficult to determine which flow features influence their predictions. This is particularly important in network management and security monitoring, where understanding model decisions can be as important as classification performance.

The analysis shows that these challenges are closely interconnected. Traffic diversity, evolving applications, and encryption increase the need for representative datasets and informative flow features, while changes in traffic characteristics can reduce model generalization. More complex models may better capture these variations, but often require greater computational resources and provide lower interpretability. Network traffic classification in modern telecommunication networks is no longer limited to developing models that achieve high classification accuracy. For practical deployment, a classifier should be able to adapt to evolving traffic patterns, handle new or previously

unseen classes, process large volumes of traffic under limited computational resources, and provide interpretable decisions. Therefore, the development and evaluation of future traffic classification systems should consider multiple criteria, including classification performance, generalization ability, adaptability, computational efficiency, and model interpretability.

References

1. Shi-Jie Xu, Kai-Chun Kong, Xiao-Bo Jin, Guang-Gang Geng, “Unveiling traffic paths: Explainable path signature feature-based encrypted traffic classification”. *Computers & Security*, Vol. 150, 2025, Article 104283. DOI: 10.1016/j.cose.2024.104283.
2. A. Azab, M. Khasawneh, S. Alrabaee, K.-K. R. Choo, and M. Sarsour, “Network traffic classification: Techniques, datasets, and challenges” *Digital Communications and Networks*, vol. 10, no. 3, pp. 676–692, 2024, doi: 10.1016/j.dcan.2022.09.009.
3. Z. Li, L. Bu, Y. Wang, Q. Ma, L. Tan, and F. Bu, “Hierarchical perception for encrypted traffic classification via class incremental learning”. *Computers & Security*, vol. 149, Art. no. 104195, 2025, doi: 10.1016/j.cose.2024.104195.