Internet Engineering Task Force
Internet-Draft
Intended status: Experimental
Expires: 20 September 2026

H. Yu
Huazhong University of Science and Technology

19 March 2026

Network Traffic Analysis and Network Modal Mapping Method

draft-traffic-analysis-and-network-mode-mapping-02

Abstract

This document presents a framework for network traffic classification and modality mapping based on large language models (LLMs), addressing the inefficiencies of traditional methods in dynamic network environments. The proposed approach automates multi-dimensional traffic feature extraction and intelligent decision-making to achieve precise alignment between traffic patterns and computing-storage-transmission requirements. The framework comprises two phases: pre-training (generating multi-modal traffic representations from pcap data) and mapping (dynamically formulating resource allocation strategies). It supports anomaly detection, QoS assurance, and multi-service collaboration, thereby significantly enhancing resource utilization efficiency and network service performance.

Status of This Memo

Copyright Notice

Table of Contents

1. Status of This Memo

2. Copyright Notice

3. Introduction

This document presents a novel framework that employs Large Language Models (LLMs) to automate network traffic classification and resource mapping. As network traffic experiences exponential growth, infrastructure complexity increases, and vertical industries exhibit converged requirements for storage, transmission, and computing resources in intelligent computing applications, traditional approaches relying on manual feature engineering and static classification methods have become inefficient and inadequate for dynamic network environments.

The proposed framework addresses these challenges through the following methodology: Raw pcap traffic data is first transformed into multi-dimensional feature representations. These features are then processed by LLMs to generate adaptive traffic classification models capable of recognizing diverse network flow patterns. Subsequently, the framework performs intelligent resource allocation by mapping classified traffic types to corresponding network requirements (e.g., bandwidth, latency, reliability). Finally, the framework establishes accurate and dynamic mappings between network traffic patterns and their corresponding resource requirements through continuous learning and optimization mechanisms.

4. Scope

This framework applies to network operators and service providers requiring dynamic Quality of Service(QoS) management, anomaly detection, and application-aware resource allocation. It defines methodologies for integrating LLMs into traffic analysis pipelines and mapping multi-dimensional SFC features to network modalities.

4.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

5. Terms and Definitions

Pre-training: Pre-training refers to the initial training process of a model on large-scale unsupervised data, aiming to learn general features and patterns within the data. This stage typically adopts self-supervised learning methods, such as masked language modeling (MLM) or autoregressive language modeling (AR).

Fine-tuning: Fine-tuning involves training a pre-trained model on supervised data specific to a particular task, allowing it to adapt to specialized datasets and improve performance for specific applications.

Pcap: A file format used to store network traffic data for analysis. Pcap files record packets transmitted in a network, including source and destination addresses, protocol types, payload content, and other key information.

CSV: A file format used for storing structured tabular data. CSV files use commas to separate fields, making them suitable for data analysis, machine learning training datasets, and information exchange between different applications.

Large Language Model-A deep learning model with billions or even trillions of parameters capable of processing and generating natural language text. LLMs are trained on large-scale text datasets and possess abilities such as comprehension, summarization, translation, and reasoning. They are widely used in question-answering systems, dialogue generation, text summarization, and other natural language processing tasks.

Multi-level Flow Representation-A method for representing and processing network traffic data by extracting features at multiple levels, such as packet level, flow level, and session level. MFR enables a more comprehensive analysis of network traffic characteristics.

Low-Rank Adaptation-An efficient fine-tuning method for LLMs that introduces low-rank matrices in the parameter space of a pre-trained model. LoRA reduces computational and storage costs while maintaining model performance.

Quality of Service-A network mechanism that ensures performance metrics such as throughput, latency, jitter, and packet loss rate during data transmission. QoS is critical for applications that require stable and predictable network performance.

Guaranteed Bit Rate-A QoS mechanism that ensures a minimum bit rate for a specific data flow. GBR is used in applications requiring stable bandwidth allocation, such as voice calls and video conferencing, to maintain consistent service quality.

Non-Guaranteed Bit Rate-A QoS mechanism where no minimum bit rate is guaranteed for data flows. Non-GBR is suitable for applications with flexible bandwidth needs, such as web browsing and social media.

5G QoS Identifier-A QoS identifier in 5G networks used to differentiate various types of traffic and define QoS requirements for different services. 5QI helps ensure appropriate network resource allocation for applications like high-definition video streaming, cloud gaming, and industrial automation.

6. Abbreviations

LLM: Large Language Model

MFR: Multi-level Flow Representation

LoRA: Low-Rank Adaptation

Qos: Quality of Service

GBR: Guaranteed Bit Rate

Non-GBR: Non-Guaranteed Bit Rate

5QI: 5G QoS Identifier

7. Framework Overview

The framework operates in two phases:

1. Pre-training Phase: Converts raw pcap data into byte streams, CSV features, MFR matrices, and traffic graphs. Fine-tunes an LLM using LoRA to generate a traffic classification model.

2. Demand Mapping Phase: Applies the model to classify live traffic and maps categories to network modalities via predefined rules. Continuously optimizes configurations based on real-time feedback.

```
+-------------+   +-------------------------+  +---------------------+    +--------------
------------+
| Raw Traffic | --> | Byte Stream CSV Features | --> | Large Language Model |  --> | Traf
fic Classification Model |
| Data     |        |  MFR Matrix Traffic Graph |   | (Low-Rank Adaptation) |     +-------
-------------------+
+-------------+  +-------------------------+  +---------------------+
          |                             |
          V                            V
+-----------------------------+          +---------------------+
| Pre-collected Knowledge Base |          | Modal Network Demand |
| Traffic Characteristics      |     -->  | Mapping Rules        |
+-----------------------------+          +---------------------+
                        |
                        V
+------------------+  +------------------+    +-----------------------+  +--------------
-----------+
| Demand Mapping  | <--| Traffic Classification |    | Byte Stream, CSV Features | <-- |
Raw  Traffic Data |
| Configuration   |   |   Model                |    | MFR Matrix Traffic Graph  |   +----
---------------------+
+------------------+  +------------------+    +------------------------+
                                                          |
                                                          V
+----------------------+     +---------------------------+     +----------------------+
| Based on User Demand  |      | Traffic Analysis Storage |     | Network Verification |
| Estimation            |  +   | Transmission Computation |     +----------------------+
+----------------------+     +----------------------------+
```

Figure 1

8. Use Cases

8.1. Dynamic Resource Allocation

Prioritizes high-QoS traffic (e.g., video conferencing) by allocating guaranteed bandwidth via GBR, while deprioritizing Non-GBR traffic (e.g., file downloads).

8.2. Anomaly Detection

Identifies deviations from learned traffic patterns (e.g., DDoS attacks) and triggers alerts for mitigation.

8.3. Application-Aware Networking

Maps application-specific traffic (e.g., streaming video) to predefined 5QI profiles for optimized QoS.

9. Security Considerations

Implementations MUST anonymize pcap data to prevent leakage of sensitive information. LLM models and knowledge bases SHOULD be protected against unauthorized access. Real-time monitoring is RECOMMENDED to detect adversarial inputs.

10. IANA Considerations

This document makes no requests to IANA.

11. References

11.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", RFC 8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

Authors’ Addresses

Hui Yu Huazhong University of Science and Technology Wuhan, China
Email: hui_yu@hust.edu.cn

Author’s Address

Hui Yu
Huazhong University of Science and Technology
Wuhan
China
Email: hui_yu@hust.edu.cn