Internship / Ph.D. Thesis Proposal 2024-2025

Title: Generative AI Techniques for Network Management

Host laboratory: LIP, ENS de Lyon, 46 allée d'Italie, Lyon, France

Advisors:

Francesco Bronzino MCF HDR ENS Lyon `francesco.bronzino@ens-lyon.fr`

External Collaborators: The thesis will be carried on within a collaborative project with the University of Chicago and ICSI at UC Berkeley.

Starting Date: The thesis will start in October 2025 and will be preceded by a Master internship.

Keywords: Traffic analysis, Generative AI, Network Management, Machine Learning.

Description.

Inferring the quality of applications and services is important to Internet Service Providers (ISPs) to detect network issues and facilitate capacity planning. However, with the widespread adoption of end-to-end encryption, ISPs cannot directly observe application quality metrics directly from the traffic traversing their network. Despite traffic becoming more opaque, recent work has shown that it is indeed possible to infer quality metrics by training a machine learning model on traffic features. And yet, these models have not made the transition to practice at ISPs. While inference models have shown to be accurate, their adoption has been slowed by factors that go beyond model design. One key challenge to achieve good models accuracy in practice is the need for high similarity between the data used at training time and the one observed in deployment. However, labeled data is challenging to collect due to high collection costs, privacy concerns, and limited access to application ground truth. Enlarging real-life datasets with high-fidelity synthetic data has emerged as a remedy for the data-scarcity problem. In the recent past, Generative Adversarial Networks (GANs) have been the conventional technique used for generating realistic network traffic. However, training GANs for network traffic requires careful consideration of data distribution and model interpretability. Further, GAN-based methods tend to focus on a limited set of attributes or statistics, are notoriously hard to train, and show limited long term generation stability.

In this project we will study the role of generative AI techniques in solving the unanswered problems of scarcity of data to train quality inference models and translate inference models' output into actions. We plan to tap into the unique advantages offered by generative AI models to develop new tools aimed at deploying end-to-end quality inference models, namely generating synthetic network traces to train models on. Very recent approaches based on pretrained foundation models have shown promise in generating more realistic packet-level traces. In particular, our recent work has show that controlled text-to-image diffusion models are a viable solution to generate synthetic raw network traffic that complies with transport and network layer protocol rules [1]. Yet, several limitations remain open for exploration, especially when targeting quality inference ML models. First, diffusion models depend on a constant image size for both training and generation, constraining synthetic traces to a fixed length. Second, while diffusion models are highly expressive, resulting in synthetic traces that more accurately mimic real network dynamics, they produce noisy outputs that can compromise the correctness of generated traces and do not account for inter packet-timings. Third, diffusion models are incapable of capturing complex correlations between traces and their associated metadata, i.e., the application quality associated with a network trace. To address these challenges, we aim to explore generative AI techniques that are more suited for the generation of complex sequential data. For example, transformers have shown efficacy in generating sequential data like text, suggesting their potential for network traffic generation. Key challenges to achieve this goal include appropriate packet capture tokenization and maintaining long contexts for generating meaningful flows, generating semantically meaningful traffic "payloads" (i.e., data within each packet), and generate multi-dimensional time series of collection of flows and associated metadata.

Candidate Requirements.

- The candidate should have completed a qualifying program by the starting date of the thesis.
- Comfortable speaking English or French (French is not required).
- Good understanding of computer networks protocols and machine learning.
- Good proficiency with at least one programming language, preferably Golang or Rust.

What to submit. An up to date CV, university transcripts, and a letter of motivation clearly stating what the motivations to work on the described subject. One or (preferably) two recommendation letters are also welcome and strongly encouraged.

References

[1] X. Jiang, S. Liu, A. Gember-Jacobson, A. Nitin Bhagoji, P. Schmitt, F. Bronzino, and N. Feamster. Netdiffusion: Network data augmentation through protocol-constrained traffic generation. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2024.