<div align="center">**Internship Proposal 2025-2026**</div>

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
Adrien Schoen	Postdoc ENS Lyon	adrien.schoen@ens-lyon.fr

Starting Date: The internship will be done during spring / summer 2026. It could be followed by a Ph.D. thesis. The thesis could start in October 2026 and will be preceded by a Master internship.

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

Description. Internet Service Providers (ISPs) need to infer application quality metrics from encrypted traffic to detect network issues and plan capacity. While machine learning models can infer quality from traffic features, their deployment requires large labeled datasets that are expensive to collect and raise privacy concerns. Generating synthetic network traffic offers a promising solution to this data scarcity problem. Our recent work on NetDiffusion [1] demonstrated that controlled text-to-image diffusion models can generate realistic packet-level traces that comply with protocol rules. However, several limitations prevent their use for training quality inference models: (1) fixed-length trace generation due to constant image size requirements, (2) noisy outputs that may violate protocol correctness and lack inter-packet timing information, and (3) inability to capture correlations between traces and their associated quality metrics.

This internship will address these limitations by exploring generative AI techniques better suited for sequential network data. The student will investigate transformer-based approaches, which have proven effective for sequential data like text, and evaluate their potential for network traffic generation. Key technical challenges include designing appropriate tokenization schemes for packet captures, maintaining long contexts for meaningful flow generation, generating semantically valid packet payloads, and conditioning generation on quality metadata.

Deliverables. The intern will implement and evaluate a prototype traffic generator, compare it against NetDiffusion baselines, and assess whether synthetic traces improve quality inference model accuracy. The work will potentially result in a research paper submission.

Candidate Requirements.
- The candidate should be enrolled in a qualifying program by the starting date of the internship.
- Comfortable speaking English or French (<u>French is not required</u>).
- Good understanding of computer networks protocols and machine learning.

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.