Chao WANG

Location: Paris, France
Email: chao.wang@eurecom.fr | Mobile: +330760457779/+8618519641239
Personal Website: https://chao0511.github.io/chaowang-website/

Summary

I am a third-year Ph.D. candidate (CIFRE) at EURECOM, Sorbonne university and HUAWEI Research Center, advised by Pietro MICHIARDI, Alessandro FINAMORE, and Giulio Franzese. Before that, I received my double master/engineering degree in Data Science from CentraleSupélec, Paris-Saclay University and in System Engineering from Beihang University, and my bachelor's degree in Mathematics and Applied Mathematics from Beihang University. I also interned at HUAWEI Paris Research Center and EDF R&D France. My research focuses on generative models for multimodality and data augmentation for network Traffic Classification.

Education

EURECOM, Sorbonne University & HUAWEI Research Center — Paris, France
Ph.D student (CIFRE) in Computer Science, department of DataCom-AI4Net — *Sep 2022 – Sep 2025*

- Mentors: Prof. Pietro MICHIARDI, Alessandro FINAMORE, Giulio FRANZESE

CentraleSupélec, Paris-Saclay University — Saclay, France
M.E. (Diplôme d'ingénieur) in Data Science and Information, GPA 4.1/4.3 — *Sep 2018 – Dec 2021*

- Computer Science Courses: Deep learning, Natural language processing, Computer vision, Reinforcement learning, Graph neural network, Machine Learning
- Mathematical Courses: Harmonic analysis, Stochastic process, Distributed optimization, Optimization, Metric theory, Statistics, Partial differential equation

Beihang University (formerly as Beijing University of Aeronautics and Astronautics) — Beijing, China
M.E. in System Engineering, double degree — *Sep 2019 – Jun 2022*

Beihang University — Beijing, China
B.S. in Mathematics and Applied Mathematics — *Sep 2015 – Jun 2019*

Work Experience

HUAWEI Research Center — Paris, France
DataCom - AI4Net Group — *Sep 2022 – Sep 2025*

- Mentors: Alessandro FINAMORE, Massimo GALLO
- As a Cifre Ph.D. student, my research focuses on:
  * Generative models for multimodality: Generative models (e.g. Diffusion Models, Flow Matching models) for multimodal conditional generation (e.g. Text-to-Image) still face challenges on aligning the generated samples with complex condition signals. I develop generative models-based Multual Information (MI) estimators, and use MI to guide model alignment.
  * Data augmentation for network Traffic Classification (TC): In TC, large-scale datasets with high-quality labeling is expensive. To alleviate this burden, I develop augmentation methods by hand-crafted functions and generative models to boost classifiers' generalization and robustness in multiple use cases.

HUAWEI Research Center — Paris, France
DataCom - AI4Net Group — *Apr 2021 – Oct 2021*

- Mentors: Alessandro FINAMORE, Lixuan YANG, Massimo GALLO, Dario ROSSI
- I studied the problem of real-time network traffic flow classification. I proposed a lightweight ShuffleNet-based neural network model, taking TCP or UDP traffic flow as input, to predict among 200 applications. I implemented the model with Pytorch, Tensorflow and Tensorflow-Lite, and tested its performance on GPU, CPU x86 and CPU ARM. Compared with CNN, MobileNet and other baselines, the proposed model achieved a 90% classification accuracy and a speed increase of 6 to 8 times.

EDF R&D Saclay, France

Electric Mobility Group *Jul 2020 – Sep 2020*

- Mentor: Elliot BUTTY
- I studied the problem of Electric Vehicle (EV) charging prediction. Given the EV charging data in Beijing, I implemented various machine learning algorithms for behavioral analysis and modelling, including xgboost, LSTM, clustering, etc. It serves to allocate electricity more scientifically.

Publications and Prinprints

RFMI: Estimating Mutual Information on Rectified Flow for Text-to-Image Alignment

- In International Conference on Learning Representations - Workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy (ICLR Workshop)'25
- Chao Wang, Giulio Franzese, Alessandro Finamore, Massimo Gallo, Pietro Michiardi

Information Theoretic Text-to-Image Alignment

- In International Conference on Learning Representations(ICLR)'25
- Chao Wang, Giulio Franzese, Alessandro Finamore, Massimo Gallo, Pietro Michiardi

Video-Based Recognition of Online Learning Behaviors Using Attention Mechanisms

- In IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)'24
- Bingchao Huang, Chuantao Yin, Chao Wang, Hui Chen, Yanmei Chai, Yuanxin Ouyang

Data Augmentation for Traffic Classification

- In Passive and Active Measurement Conference (PAM)'24, *Runner-up Best Paper Award*
- Chao Wang, Alessandro Finamore, Pietro Michiardi, Massimo Gallo, Dario Rossi

Toward Generative Data Augmentation for Traffic Classification

- In Conference on emerging Networking EXperiments and Technologies (CoNEXT)'23, Student Workshop
- Chao Wang, Alessandro Finamore, Pietro Michiardi, Massimo Gallo, Dario Rossi

Contrastive Learning and Data Augmentation in Traffic Classification Using a Flowpic Input Representation

- In Proceedings of ACM on Internet Measurement Conference (IMC)'23
- Alessandro Finamore, Chao Wang, Jonatan Krolikowski, Jose M Navarro, Fuxing Chen, Dario Rossi

Many or Few Samples? Comparing Transfer, Contrastive and Meta-Learning in Encrypted Traffic Classification

- In Network Traffic Measurement and Analysis Conference (TMA)'23
- Idio Guarino, Chao Wang, Alessandro Finamore, Antonio Pescape, Dario Rossi

AppClassNet: a commercial-grade dataset for application identification research

- In ACM SIGCOMM Computer Communication Review'22
- Chao Wang, Alessandro Finamore, Lixuan Yang, Kevin Fauvel, Dario Rossi

Academic Service

Reviewing for Conferences : SIGCOMM'24 Artifacts Evaluation

Reviewing for Journals : Computer Networks

Key Skills

Languages : French, English, Chinese

Programming Languages : Python

Machine Learning Libraries : PyTorch, Tensorflow, Scikit-Learn, NumPy, Pandas