Fabian Spaeh

Contact no. +1 (857) 529 5666
E-mail fspaeh@bu.edu
Website 285714.github.io

Education

Ph.D. in Computer Science (GPA 3.9) Sep 2020 – May 2025
Boston University, Advised by Prof. Alina Ene Boston, MA

- Thesis "Efficient Algorithms for Online Resource Allocation and Submodular Maximization"
- NSF Travel Grant in 2024
- Teaching Fellow Excellence Award in 2022
- Dean's Fellowship in spring 2021
- Coursework included Advanced Optimization Alg., Advanced Topics in CS Graph Analytics, Taming Big Data

M.Sc. in Computer and Information Science (VEUK award, GPA 4.0) Apr 2018 – Apr 2020
University of Konstanz Konstanz, Germany

- Coursework included Methods of Network Analysis, Randomized Algorithms, Mathematics for Data Science

B.Sc. in Computer Science (VEUK award for academic excellence, GPA 3.9) Oct 2013 – Feb 2018
University of Konstanz Konstanz, Germany

Research

Learning-Augmented Optimization Algorithms
Using machine learning to improve the performance of algorithms that act under uncertainty

- Designed the first algorithms with predictions for general online assignment problems [8, 7]. This is relevant for ad allocation, submodular maximization, and portfolio optimization.
- Currently working on learning-augmented algorithms for resource allocation that build upon multiplicative weights-based procedures and optimal transport.

Machine Learning
Designing and analyzing machine learning models; devising and testing algorithms for efficient learning

- Developed the first methods to learn mixtures of continuous-time Markov chains, with applications to sports analytics and time-series analysis [6, 5]. Predictive performance surpasses state-of-the-art deep learning models.
- Generalized a method for efficiently learning discrete-time Markov chain mixtures to important topologies [9].

Work Experience

Applied Scientist Jun 2025
Celonis AI New York City, NY
Intern, Machine Learning Jun – Aug 2024
Celonis AI Palo Alto, CA

- Query Suggestion for Retrieval-Augmented Generation (RAG): Developed a novel self-learning method to recommend answerable queries to unexecutable user queries, with over 93% answerability on multi-step queries.
- Deployed the solution within the Celonis Copilot (Python).

Intern, Quantitative Researcher Jun – Aug 2023
TWT, Mathematics, Computer Graphics & Sustainability Engineering Munich, Germany

- Cable Harness Optimization: Created an efficient method to optimize cable harnesses design under complex constraints, balancing harness weight and assembly complexity by extending algorithms for Steiner trees.
- Supervised a graduate student in application development (C#).

Intern, Data Science Jun – Aug 2020
German Federal Bank (Eurosystem), Division Monetary and Financial Statistics Frankfurt, Germany

- Data Cleansing and Imputation: Aggregated data and trained models for missing value imputation in the Register of Institutions and Affiliates Data for financial reporting (Python: Pandas, Dask, scikit-learn).
- Imputed 300 000 values with a precision over 96%. Automatically identified and flagged data quality issues.

Student Research Assistant Jan – Oct 2018
University of Konstanz Konstanz, Germany

- Creation of soccer live-betting data: Distributed collection of a large dataset with 270 000 000 betting odds.
- Learning and execution of betting strategies (Python, SQL, Redis).

Teaching

Boston University
Teaching Assistant

- Randomness in Computing, Graduate Class. Fall 2021 and Fall 2022
- Advanced Optimization Algorithms, Graduate Class. Spring 2022 and Fall 2023

University of Konstanz
Lab Instructor for Analysis and Linear Algebra, Discrete Mathematics and Logic, and Programming Course 2

Technical Skills

Proficient in Python (PyTorch, scikit-learn, SciPy), mathematical programming (GLPK, Gurobi, cvxpy), C#, C++, Java, JavaScript, SQL, and Haskell.

Publications and Manuscripts

[1] F. Spaeh and A. Miyauchi, "An asymptotically optimal approximation algorithm for multiobjective submodular maximization at scale." ICML 2025.

[2] F. Spaeh, T. Chen, C.-H. Chiang, B. Shen, and C. Yu, "Query suggestion for retrieval-augmented generation via dynamic few-shot learning at celonis." In submission.

[3] D. Ristache, F. Spaeh, and C. Tsourakakis, "Countering election sway: Strategic algorithms in friedkin-johnsen dynamics." In submission.

[4] D. Ristache, F. Spaeh, and C. Tsourakakis, "Wiser than the wisest of crowds: The Asch effect and polarization revisited." ECML PKDD 2024.

[5] F. Spaeh, K. Sotiropoulos, and C. Tsourakakis, "ULTRA-MC: A unified approach to learning mixtures of markov chains via hitting times." In submission.

[6] F. Spaeh and C. Tsourakakis, "Markovletics: Methods and a novel application for learning continuous-time markov chain mixtures." WWW 2024.

[7] F. Spaeh, A. Ene, and H. L. Nguyen, "Online and streaming algorithms for constrained k-submodular maximization." AAAI 2025.

[8] F. Spaeh and A. Ene, "Online ad allocation with predictions." NeurIPS 2023.

[9] F. Spaeh and C. Tsourakakis, "Learning mixtures of markov chains with quality guarantees." WWW 2023.

[10] F. Spaeh and S. Kosub, "Global evaluation for decision tree learning." arXiv, 2022.

[11] T. Hepp, F. Spaeh, A. Schönhals, P. Ehret, and B. Gipp, "Exploring potentials and challenges of blockchain-based public key infrastructures." IEEE INFOCOM Workshops, 2019.