Kelly Kostopoulou

(+1) 646-286-9074 | kelkost@cs.columbia.edu | delphiancalamity.github.io | DelphianCalamity

EDUCATION

Columbia University, New York, USA

- **PhD in Computer Science** — Sep 2020 - Present
  - Current GPA: 4.22/4.00, Advisor: Asaf Cidon

National & Kapodistrian University of Athens (NKUA), Athens, Greece

- **MSc in Data Science & Information Technologies** — Sep 2018 - Sep 2020
  - GPA: 9.7/10 *(first among graduate class of 10 students)*
  - *Specialization: Big Data and Artificial Intelligence*
  - *Thesis: Sparse Communication for Deep Learning*
- **BSc in Informatics & Telecommunications** — Oct 2012 - Sep 2018
  - GPA: 9.01/10 *(first among graduate class of 79 students)*
  - *Specialization: Data and Knowledge Management | Software*
  - *Thesis: Recursive Function Definitions in Static Dataflow Graphs & their Implementation in TensorFlow*

HONORS & AWARDS

2024 Avanessians Fellowship — Jan 2024

Gerondelis Foundation Scholarship — May 2023

Bodossakis Foundation Scholarship — Sep 2022

Onassis Foundation Scholarship — Sep 2022

Valedictorian, MSc, Dept. of Informatics and Telecommunications, NKUA — Jan 2021

Valedictorian, BSc, Dept. of Informatics and Telecommunications, NKUA — Mar 2019

Scholarship of Excellence, MSc of Data Science & Information Technologies, NKUA — Oct 2018

EXPERIENCE & RESEARCH

Gray Systems Lab, Microsoft Research, Redmond, Washington — Jun 2024 - Aug 2024

Research Intern

Visiting Student Research Program, KAUST, Saudi Arabia — Jan 2020 - Jun 2020

Research Student

- I developed a compressed communication mechanism for distributed deep learning using bloom-filters.

National Center of Scientific Research - Demokritos, Athens, Greece — Sep 2018 - Dec 2019

Research Software Engineer, Research Assistant

- I participated in the implementation of the National Network of Precision Medicine for Cancer Prevention and Treatment where I created a cloud-based platform for the distributed execution of reproducible scientific workflows, developed bioinformatic pipelines and created a web application for user interaction.

Dab Ltd, Athens, Greece — Sep 2016 - Sep 2018

IT Consultant

PUBLICATIONS

- Pierre Tholoniat*, Kelly Kostopoulou*, Peter McNeely, Prabhpreet Singh Sodhi, Anirudh Varanasi, Benjamin Case, Asaf Cidon, Roxana Geambasu, Mathias Lecuyer. "Cookie Monster: Efficient On-device Budgeting for Differentially-Private Ad-Measurement Systems", SOSP '24 (*equal contribution)
- Kelly Kostopoulou*, Pierre Tholoniat*, Asaf Cidon, Roxana Geambasu, Mathias Lecuyer. "Turbo: Effective caching in Differentially-Private Databases", SOSP '23 (*equal contribution)
- Pierre Tholoniat*, Kelly Kostopoulou*, Asaf Cidon, Mosharaf Crowdhury, Roxana Geambasu, Mathias Lecuyer, Junfeng Yang. " Efficiently Packing Privacy Budget with DPack", EuroSys '25 (*equal contribution)
- Hang Xu, Kelly Kostopoulou, Aritra Dutta, Xin Li, Alexandros Ntoulas, Panos Kalnis,'DeepReduce: A Sparse-tensor Communication Framework for Federated Deep Learning", NeurIPS '21

ADDITIONAL INFORMATION

Technical skills: C/C++, Java, Python, SQL, TensorFlow, Docker, Kubernetes, Git
Languages: **English** (proficient), **French** (basic), **Greek** (native)

RESEARCH PORTFOLIO

Cookie Monster: Efficient On-device Budgeting for Differentially-Private Ad-Measurement Systems

Cookie Monster is an on-device budgeting component that can be integrated into differentially private ad-measurement systems. Powered by a robust theoretical framework known as Individual Differential Privacy (IDP), a variant of traditional differential privacy, Cookie Monster allows advertisers to conserve significantly more privacy budget compared to existing alternatives.

Publication: Pierre Tholoniat*, Kelly Kostopoulou*, Peter McNeely, Prabhpreet Singh Sodhi, Anirudh Varanasi, Benjamin Case, Asaf Cidon, Roxana Geambasu, Mathias Lecuyer. "Cookie Monster: Efficient On-device Budgeting for Differentially-Private Ad-Measurement Systems", SOSP '24 (*equal contribution)

Turbo: Effective caching in Differentially-Private Databases

Effective caching for linear query workloads over DP databases. Turbo builds upon private multiplicative weights (PMW), a DP mechanism that is powerful in theory but very ineffective in practice, and transforms it into a highly effective caching object, namely PMW-Bypass. A description of this project can be found on our paper, titled Turbo: Effective caching in Differentially-Private Databases.

Publication: Kelly Kostopoulou*, Pierre Tholoniat*, Asaf Cidon, Roxana Geambasu, Mathias Lecuyer. "Turbo: Effective caching in Differentially-Private Databases", SOSP '23 (*equal contribution)

DPack: Efficiently Packing Privacy Budget

DPack is a framework that aims to maximize the exploitation of private datasets when using differential privacy. Differential privacy theoretically bounds the privacy loss when querying a dataset. Every time a data scientist analyzes a dataset through the execution of queries they "consume" privacy budget. Since privacy loss is bounded there is only a finite number of queries that a dataset can answer before breaking the privacy guarantees. In this work, we view the privacy that will be consumed by the queries as a new scheduling resource, which is a first-class citizen alongside other traditional computing resources, such as CPU, memory and network bandwidth. The goal of DPack is to execute queries in a schedule that maximizes the total number of queries that can be processed or the total utility that can be achieved given also the importance of each query. In this work, we reduce this scheduling problem to an optimization problem and implement a tool that simulates the dynamic arrival of data and queries and a scheduler that is responsible for allocating privacy resources for the execution of the queries. Finally, we incorporated the scheduler to a real, open-sourced resource orchestrator to enable its broader applicability. We chose to integrate DPack with Kubernetes, the most popular resource orchestrator in cloud systems. I took part in all aspects of this project, and was solely responsible for the systems side: the design and implementation of DPack in Kubernetes.

Publication: Pierre Tholoniat*, Kelly Kostopoulou*, Asaf Cidon, Mosharaf Crowdhury, Roxana Geambasu, Mathias Lecuyer, Junfeng Yang. " Efficiently Packing Privacy Budget with DPack", EuroSys '25 (*equal contribution)

Sparse Communication for Deep Learning

I led this project as part of my MSc thesis at KAUST. I developed a compressed communication mechanism for distributed machine learning using bloom filters. When training neural networks in a distributed setting with data parallelism each replica of the model generates its own local gradients. Those gradients need to be broadcasted and aggregated so that they are used for the update of the replicas. This creates a significant communication bottleneck that gets mitigated through the compression of the gradients. I developed two TensorFlow operators for encoding and decoding the indices of sparsified gradients, as those were produced during the training of deep neural networks. The encoded output was a bloom-filter which is lossy in nature, so the inaccurate reconstruction of the gradients was degrading the accuracy of the models. To remedy that, I developed the "false-positive-aware" mechanism, a compression method that adaptively modifies the encoded gradient based on its knowledge for all the future bloom-filter errors.

Thesis: Kelly Kostopoulou, "Sparse Communication in Deep Learning"

Publication: Hang Xu, Kelly Kostopoulou, Aritra Dutta, Xin Li, Alexandros Ntoulas, Panos Kalnis, "DeepReduce: A Sparse-tensor Communication Framework for Federated Deep Learning", NIPS21

Recursive Function Definitions in Static Dataflow Graphs & their Implementation in TensorFlow

I led this project as part of my BSc thesis at University of Athens. I investigated how recursive functions can be supported in dataflow systems and especially TensorFlow. The challenge was to create dataflow graphs that can capture the logic of recursive functions without being dynamically expanded at runtime. Static dataflow graphs enable a variety of graph optimizations that can accelerate the training of recursive neural networks, which are widely used in natural language processing. I enabled the support of recursion in TensorFlow through extending its API, enforcing graph transformations that make the graphs static during their execution and extending its execution engine with a tagging mechanism that enables the correct execution of those graphs.

Thesis: Kelly Kostopoulou, "Recursive Function Definitions in Static Dataflow Graphs & their Implementation in TensorFlow"

Talk: 16th Programming Languages Seminar, Dec 2018, National and Technical University of Athens (NTUA)