Andy Huynh

PhD Candidate in Computer Science

huynh.n.andy@gmail.com · ndhuynh.com · United States

Research Interest

Databases, Tuning and Optimizing Data Systems, Robust Data Systems, High Performance Computing

Education

Ph.D. Computer Science 2017 - Present
Boston University with Manos Athanassoulis

B.Eng Computer Engineering *Magna Cum Laude* with *Distinction* 2014 - 2017
University of Minnesota, Twin Cities with Vladimir Cherkassky

Awards

IBM Ph.D. Student Fellowship *2020*

Dean's List 6 Semester at University of Minnesota *2014 - 2017*

Experience

Ph.D. Researcher at Boston University with Manos Athanassoulis 2017 - Present

Research Intern at Netapp in Advanced Technology Group 2019

Machine Learning Research Intern at Bose in Automotive Group 2018

Undergraduate Researcher at University of Minnesota with Vladimir Cherkassky 2016 - 2017

Firmware Engineering Intern at Medtronic 2017

Undergraduate Researcher at University of Minnesota with Dr. Bin He 2015 - 2017

Undergraduate Researcher at University of Minnesota with Jack Stubbs, Fluvio Lobo 2014 - 2017

Publications

1. *Endure: A Robust Tuning Paradigm for LSM Trees Under Workload Uncertainty*
   Huynh A, Chaudhari H, Terzi E, Athanassoulis M. *Under Submission*

2. *Analysis of Allogeneic Hematopoietic Cell Transplant Patients Using Self-Organizing Maps*
   Huynh A, Cherkassky V. Undergraduate Honors Thesis, 2017,
   Department of ECE, University of Minnesota Twin Cities.

3. *Modeling of Swine Diaphragmatic Tissue Under Uniaxial Loading*
   Huynh A, Molina Espinosa M, Lobo Fenoglietto F, Singal A, Iaizzo P.
   ASME Journal of Medical Devices, 9(3), 3-3. 2015.

Posters and Presentations

1. *Robust LSM-Trees: Tuning for Workload and Resources Uncertainty*
   RedHat Greater New England Research Interest Group Meeting, Presentation, May 2021

2. *Tuning Data Systems Under Uncertainty in Workload*
   North East Database Day 2020, Poster

Teaching

CS 460: Introduction to Database Systems	Fall 2020
CS 591A: Data Systems Architecture	Spring 2020
CS 591P: Object Oriented Programming in Java	Fall 2018, Spring, Summer, Fall 2019
CS 112: Introduct to Computer Science II	Spring 2018, Summer 2018
CS 111: Introduction to Computer Science I	Fall 2017
PHYS 1302W: Introductory Physics for Science and Engineering II	Spring 2015

Selected Works

Robust LSM Trees: I am implementing a new robust tuning paradigm for LSM Trees. By framing the tuning problem as an optimization problem that takes into consideration uncertainties in the input factors, we can find a design that, when deployed, is robust to changes in the expected workload and resources. I utilize Python to create an optimization framework that solves the modeled problem, then pipes the design decisions into a C++ framework that exposes tuning knobs of RocksDB. An instance of the database is deployed and we can test this on randomized or real world workloads.

Optimizing Distributed Data Placement: I worked on creating a model for optimal data placement in a distributed data system that contains a high level of separation between the data nodes and the compute nodes. In a team of two, we created a simulation of the particular system setting, formulated the data placement problem as an optimization problem, and utilized the Mosek optimization package in C++ to show the improvements in performance and the competitive computation time com- pared to various other common heuristics and strategies.

Optimal Column Stores: I am working on creating a key-value column storage engine that takes into consideration the access profile across keys in order to optimally partition columns for future workloads. In addition I am taking a look how to formulate the partitioning problem as a robust optimization problem to take into consideration shifts in the expected workload. This system is open-source and written in C++.

B.Eng Thesis: I worked on analyzing blood factors from a cohort of cancer patients who have all undergone allogeneic hematopoietic cell transplant procedures using machine learning techniques. In particular I utilized self-organizing maps as a technique to visualize and understand relationships between the outcome of patient procedures and blood factors.

Mentoring

Caterina Caravaggio: Exchange student from Universit‘a di Bologna	Summer 2020