

# MOHIT RAJPAL

mohitr@comp.nus.edu.sg

<https://scholar.google.com.sg/citations?user=qUvSFVAAAAAJ>

[openreview.net/profile?id= Mohit\\_Rajpal1](https://openreview.net/profile?id=Mohit_Rajpal1)

[linkedin.com/in/mohit-rajpai](https://linkedin.com/in/mohit-rajpai)

[mohit-rajpai.github.io](https://mohit-rajpai.github.io)

---

## EDUCATION

- Doctor of Philosophy** | *Proposed thesis: Scaling up decision-making under uncertainty* Aug. 2018 – Present  
National University of Singapore Singapore, SG  
Supervisor: Dr. Bryan Kian Hsiang Low  
Reviewer service: AAAI×5, AISTATS×5, ICLR×4, ICML×4, NeurIPS×5, UAI×4  
Interests: ML for Systems, Machine Learning Systems, Bayesian Optimization, Gaussian Processes
- Master of Science** | *Major: Computer Science, Track: Machine Learning* Sep. 2015 – Dec. 2016  
Columbia University New York City, NY, USA
- Bachelor of Science** | *Major: Computer Science* Aug. 2008 – May 2011  
University of Illinois at Urbana-Champaign Urbana-Champaign, IL, USA

---

## PUBLICATIONS AND PREPRINTS

Citations: 219

- Ultra Low-Latency Traffic Engineering for Satellite Network** Sep. 2024  
Hao Wu, Yizhan Han, **Mohit Rajpal**, Qizhen Zhang, Jingxian Wang Citations: 0  
Under submission to MobiCom 2025  
The paper presents a novel traffic engineering (TE) solution for large-scale Low-Earth-Orbit (LEO) satellite constellations. Unlike traditional TE systems designed for static wide-area networks (WANs), this approach addresses the rapidly changing topology of satellite networks, ensuring ultra-low-latency traffic allocation. The proposed framework, named SaTE, leverages a cascaded graph neural network (GNN) to compute optimal traffic allocation with millisecond latency, accelerated by GPU parallelization. Evaluation on Starlink's 4236 satellites shows a 23.5% improvement in satisfied demand, achieving a 2738x speedup with an average runtime of 17ms compared to commercial solvers.
- Hessian-aware Bayesian optimization for decision making systems** Oct. 2023  
**Mohit Rajpal**, Lac Gia Tran, Yehong Zhang, Bryan Kian Hsiang Low Citations: 0  
Transactions on Machine Learning Research  
The paper introduces an approach to optimize decision making systems addressing challenges posed by sparse reward. A compact higher-order model is used for cooperative multi-agent decision making. This high-dimensional model is optimized using Hessian-aware Bayesian optimization. Validation demonstrates the effectiveness of the proposed approach in various benchmarks.
- Pruning during training by network efficacy modeling** Mar. 2023  
**Mohit Rajpal**, Yehong Zhang, Bryan Kian Hsiang Low Citations: 5  
Springer Machine Learning Journal  
The paper introduces a novel method for early pruning of deep neural network (DNN) neurons during training to reduce computational costs while preserving model performance. The approach models the future efficacy of DNN elements in a Bayesian manner, using efficacy data collected during training to identify and prune neurons during training. Empirical evaluations demonstrate that the proposed Bayesian early pruning improves the computational efficiency of DNN training while maintaining better model performance compared to other tested pruning approaches.
- Neural networks for efficient Bayesian decoding of natural images from retinal neurons** Dec. 2017  
N. Parthasarathy, E. Batty, W. Falcon, T. Rutten, **Mohit Rajpal**, E.J. Chichilnisky, Liam Paninski Citations: 77  
Neural Information Processing Systems (NeurIPS)  
Summary: The paper introduces a novel Bayesian method for decoding natural images from retinal ganglion cell (RGC) spiking activity, utilizing artificial neural networks for fast nonlinear decoding. The decoder, trained on natural images and simulated neural responses, outperforms linear decoding and provides insights for optimizing retinal prosthesis technologies. This work suggests that the retina may offer a more accurate representation of the visual scene than previously thought.

## **Not all bytes are equal: Neural byte sieve for fuzzing**

**Mohit Rajpal**, William Blum, Rishabh Singh

arXiv Preprint

Nov. 2017

Citations: 137

Summary: This paper introduces a new approach to enhancing fuzzing, a dynamic program analysis technique for identifying software vulnerabilities. In this work deep learning architectures are trained on fuzzing data to learn valuable locations to fuzz in input files. By integrating these models into the a greybox fuzzer, significant improvements are demonstrated in terms of code coverage, unique code paths, and crash discovery across diverse input format (e.g., ELF and XML).

## **WORK EXPERIENCE**

---

### **Research Contractor**

Microsoft Research

Jul. 2017 – Jul. 2019

Redmond, WA, USA

- Continued research and validation work on greybox fuzzing architecture
- Helped see the architecture through testing and pre-integration phase
- Work during this period led to a technical preprint and US Patent

### **Research Intern**

Microsoft Research

Mar. 2017 – Jun. 2017

Redmond, WA, USA

- Developed a novel neural network augmented fuzzing architecture to discover software faults
- Evaluated LSTM based architectures for reinforcement learning to improve fuzzing efficacy

### **Research Assistant**

Columbia University Department of Statistics

May 2016 – Dec. 2016

New York City, NY, USA

- Investigated decoding of retinal neuron activity to presented source images
- Wrote linear regressors and deep neural networks to perform decoding

### **Systems and Applications Developer**

PDT Partners LLC

Jun. 2013 – Aug. 2014

New York City, NY, USA

- Implemented communication software to place orders with stock exchanges

### **Software Development Engineer**

Microsoft Corporation, Windows Division

Jul. 2011 – May 2013

Redmond, WA, USA

- Developed automatic web indexers using an asynchronous, multithreaded workflow
- Developed a highly scalable, distributed, fault-tolerant, multi-threaded file hashing utility
- Maintained support for 32-bit compatibility layer for 64-bit Windows (WoW64)
- Performed end-to-end validation of Windows 8 device and driver telemetry prior to release of Windows 8

### **Software Development Engineer Intern**

Microsoft Corporation, Windows Division

May 2010 – Aug 2010

Redmond, WA, USA

- Researched a Windows kernel module to monitor and detect bugs in Windows drivers
- Designed a plugin that enables Windows Auto Update to service offline installation on disk images

## **TEACHING EXPERIENCE**

---

### **Machine Learning**

National University of Singapore

Spring 2019, Spring 2020, Spring 2021

Singapore, SG

### **Operating Systems I**

Columbia University

Autumn 2016, Autumn 2015

New York City, NY, USA

## **UNPUBLISHED RESEARCH WORKS**

---

### **Sensitivity conjecture**

Sensitivity is the discrete analogue to the Lipschitz constant demarking a “smooth boolean function.” The Sensitivity conjecture posited a limit on the complexity and cardinality of low sensitivity boolean functions. Having low sensitivity means few neighbors of any point on the boolean hypercube differ in function value. Surveyed literature on the sensitivity conjecture over Boolean functions. Slightly refined the upper bound on the number of sensitivity-s Boolean functions and hypothesized possible avenues for further progress. The sensitivity conjecture was proved to be true in July 2019.

Spring 2016

Columbia University

### High resolution soft real-time scheduler

Spring 2009

UIUC

This work investigated and designed a high resolution soft real time (HRSRT) scheduler, a practical scheduler for tasks (computer processes or threads) on multitasking systems. The HRSRT scheduler takes advantage of more accurate high resolution timers provided by the clockevents patch set introduced in Linux kernel 2.6.17. With the HRSRT scheduler, the scheduler resolution (capability to schedule short period tasks) was increased from 1 millisecond to 200 microseconds on commodity hardware and software.

## HONORS AND AWARDS

---

### President's Graduate Fellowship

Autumn 2018

NUS

The President's Graduate Fellowship (PGF) is awarded to a small number of NUS Computer Science PhD students. The PGF provides tuition waiver, and a stipend in return for research and teaching responsibilities. The stipend is valued at approximately \$150,000.

### Course Assistant Fellowship

Autumn 2016

Columbia University

The Course Assistant (CA) Fellowship provides for a tuition waiver, and a stipend for select high performing CAs in return for teaching responsibilities. The CA fellowship is valued at approximately \$32,000.

### Course Assistantship

Autumn 2015

Columbia University

The Course Assistantship provides a stipend in return for teaching responsibilities. The assistantship is valued at approximately \$11,000.

### Edmund J. James Scholar

UIUC

The Edmund J. James Scholar program at the University of Illinois at Urbana-Champaign is dedicated to recognizing and supporting academic achievement. Emphasizing the cultivation of critical thinking, leadership, and advanced academic skills, the program fosters a community of dedicated scholars. Participants engage in a challenging curriculum tailored to their fields of study, offering opportunities for collaborative projects and contributing to the scholarly community at the university.

### Dean's List

Autumn 2008, Autumn 2009

UIUC

Overall GPA above 75%<sup>th</sup> percentile in the Grainger College of Engineering.

## PATENTS

---

### Machine learning for constrained mutation based fuzz testing

Microsoft Research

October 2019

Techniques for constrained mutation-based fuzzing are described. In this process, a machine tests a code's input file using a fuzzing algorithm through multiple runs. Each run involves mutating one or more bytes of the input file and noting which parts of the code were executed. The machine then generates a heatmap based on these results, mapping each byte to a value indicating whether the mutation caused the execution of a previously untouched portion of the code. The fuzzing algorithm is subsequently tailored based on this heatmap to optimize testing.

## COURSEWORK AND TECHNICAL SKILLS

---

**Undergraduate computer science:** Algorithms, Artificial intelligence, Numerical methods, Operating systems design, Parallel computing, Programming languages & compilers, Theory of computation

**Undergraduate mathematics:** Basic discrete mathematics, Foundations of mathematics, Intro to combinatorics, Linear programming, Multivariable calculus, Probability theory, Real analysis

**Graduate computer science:** Advanced algorithms, Introduction to computational complexity, Introduction to databases, Introduction to cryptography, Machine learning, Uncertainty modeling in AI

**Graduate mathematics:** Differentiable Manifolds, Martingales (Probability), Fourier analysis, Abstract algebra

**Programming languages, libraries, and frameworks:** Python, C, C++, C#, F#, VHDL, SQL, Fortran, x86 Assembly, LLVM Bytecode; Tensorflow 1./2., Pytorch; .NET, Win32, NTOS Kernel, Linux Kernel, LLVM Toolchain