

KAPIL AGRAWAL

 [linkedin.com/in/kapilagraawal95](https://www.linkedin.com/in/kapilagraawal95)  kapila1@uci.edu  (+1)-949-867-8733

INTERESTS

Interested in Cloud Computing and Datacenter Networking with a special focus on large-scale failure analysis and disaster management techniques for web services.

EDUCATION

University of California, Irvine

2nd year Graduate Student, Computer Science

- CGPA: 4.00
- Graduate Research Assistant, Prof. Sangeetha Abdu Jyothi
- Graduate Teaching Assistant, CS 132 Computer Networks

Jun 2026

Irvine, CA

Delhi Technological University

B.Tech (Hons.) in Mathematics and Computing Engineering

Jun 2017

Delhi, India

EXPERIENCE

Microsoft Research

Research Fellow

Bengaluru, India

Jul 2018 - Jul 2021

Indian Institute of Technology

Embedded Systems Engineer

Himachal Pradesh, India

Jun 2017 - Jul 2018

TECHNICAL SKILLS

Languages: Python, C/C++, L^AT_EX, HTML/CSS, JavaScript, SQL, Bash, PHP

Frameworks & Services: Kubernetes, Helm, Ansible, Docker, Flask, MongoDB, SQLite, Azure, Jaeger, Redis

Toolkits and Packages: TensorFlow, Selenium, Gurobi, RLLib, NetworkX, Sklearn, D3js, lifelines, plotly

PUBLICATIONS & PATENTS

Monitoring Cloud Service Unreachability at Scale

Kapil Agrawal, Viral Mehta, Sundararajan Renganathan, Sreangsu Acharyya, Venkat Padmanabhan, Chakri Kotipalli, Liting Zhao

Proceedings of IEEE INFOCOM 2021

A Comparison of Class Imbalance Techniques for Real-World Landslide Predictions

Kapil Agrawal, Yashasvi Baweja, Ritwik Saha, Shubham Agrawal, Pratik Chaturvedi, Naresh Mali, Venkata Uday Kala, Varun Dutt

IEEE Conference on Machine Learning and Data Science, 2017

Low-cost sensor-based system for landslide monitoring and alerts

Varun Dutt, Kapil Agrawal, Shubham Agrawal, Pratik Chaturvedi, Naresh Mali, Venkata Uday Kala

Patent Application PCT/IN2018/050217. Geneva, Switzerland, 16/04/2018

PROJECTS

Automating Resilience Management

Advisor: Prof. Sangeetha Abdu Jyothi

UC Irvine

Sept 2021 - Present

- * Conducted chaos tests such as node failures and network partitioning on microservice benchmarks running on Kubernetes and identified a need to reassess the resilience strategy adopted widely by cloud infrastructure.
- * Pioneered a concise representation, *Resilience Tags*, to specify resilience objectives that can allow developers to state a broad range of requirements such as minimum network bandwidth, minimum storage, and compute requirements.
- * Building a Resilience Manager that interfaces with Kubernetes and expand control operations such as shutting down/migrating non-critical microservices to allow more critical services to run. Primarily profiling Kubernetes scheduler's codebase to understand bottlenecks to enable batch scheduling.

Monitoring Last-Mile Network Unreachability for Office365

Advisors: Dr. Venkat Padmanabhan & Dr. Sreangsu Acharyya

Microsoft

Jul 2019 - Mar 2020

- * Carried out a Microsoft-scale study of first-party services and observed that users' requests were often not directed to the nearest service point causing a bump in latency consequently prompting interest to actuate the first planet-scale measurement platform, Odin, for Microsoft's 2 billion+ users.
- * Built a highly scalable ML-based time-series detection tool, NetDetector implemented in Python without library overload, that monitors availability globally atop Odin, a client-side javascript agent, that probes Outlook by mimicking users. NetDetector reduced customer complaints from over 100+ complaints per month to zero in 3 months after deployment.
- * Developed a drill-down mechanism that performed distributed search over the vast space of possible reasons to explain the root cause of failures as a one-click auto-debug feature for on-call engineers thereby reducing the time from 2+ hours down to less than a minute.

Data-driven Networking for Exchange Datacenters

Microsoft

Advisors: Dr. Venkat Padmanabhan & Dr. Sreangsu Acharyya

Jan 2020 - Jan 2021

- * Co-led a Microsoft-scale study to understand which workloads consume the most network bandwidth and found that Migration contributes the most with 40%, of which Mailbox Migration alone contributes 70%.
- * Streamlined SQL pipelines to generate daily reports on Microsoft's cold storage, CosmosDB, that processed billions of mailboxes to study access patterns on mailboxes. Used smoothing techniques such as sliding window, moving average.
- * Modeled access patterns of individual workloads as a function of the age of emails by applying the Kaplan-Meier filter implemented using python's lifelines package on a representative mailbox population and found three buggy clients that brought about bulk accesses burdening the Office Exchange Store.

SSD Cache Modelling for Outlook Mail

Microsoft

Advisors: Dr. Venkat Padmanabhan & Dr. Sreangsu Acharyya

Jan 2020 - Jul 2021

- * Partnered with Exchange Store Team to model access patterns to devise a new cache eviction policy. Developed an open gym on SQL and python to test several caching policies and analyze performance overheads.
- * Launched a monitoring tool to track access patterns of 50+ workloads by building an efficient pipeline that implemented downstream survival analysis tasks on access pattern logs for more than 300K mailboxes in minutes every day.
- * Pioneered a testbed that replays accesses on mailboxes to compare cache eviction policies. Engineered a data-driven mechanism to auto-tune the time parameter in TTL-based cache eviction policy that outperformed LRU's 60% cache hit rate with an improved cache hit rate of 95%.

Generating Large Synthetic Microservice Dependency Graphs Using Reinforcement Learning

UC Irvine

Advisor: Prof. Sangeetha Abdu Jyothi

Sept 2022 - Present

- * Analyzing microservice benchmarks to answer questions such as can the underlying graph structure be leveraged when deploying these microservices in clouds to enhance performance and resilience yet keep costs low.
- * Investigating several toy architectures revealed that existing benchmarks fail to capture prevalent characteristics in production workloads such as microservice call invocations are highly dynamic, span 10s of services in each tier, etc.
- * Building an RL-based microservice dependency graph generator to emulate production workloads that can act as a testbed to run and compare several algorithms that arise in distributed settings such as planning, scheduling, etc.

Focused Extraction from Structured Data

Microsoft

Advisor: Dr. Arun Iyer

Jul 2018 - Dec 2018

- * Designed a highly accurate ML-based feature for Outlook that extracts information from organized content such as order summaries, flight itineraries, etc. This feature allowed Outlook to extract information on unseen templates which was previously not possible.
- * Curated a large data repository with >10000 labeled examples as a testbed to examine ML models on information retrieval tasks such as email summarization, and autofill feature in Excel with Microsoft-wide adoption.
- * Devised a technique that leverages visual cues in the structured text to form segmentation trees to generate embeddings for entire text blocks. Assembled a semi-supervised classifier which improved the overall precision of the two-staged ML model from 90% to 99% with <100 data points.