

Caleb J. Wang

cjwang.io | caleb.wang@northwestern.edu | [theheokboi](https://github.com/theheokboi) | [in hyes92121](https://www.linkedin.com/in/hyes92121)

EDUCATION

Northwestern University

Ph.D. Candidate in Computer Science

Advised by Prof. [Fabian E. Bustamante](#)

Affiliated with [AquaLab](#)

Evanston, IL, USA

Sep. 2022 – Present

National Taiwan University

Master of Science in Communication Engineering

Taipei, Taiwan

Feb. 2020 – Jun. 2021

Thesis: *Discovering Twitch's Video Delivery Infrastructure Utilizing Cloud Services and VPNs*

Advised by Prof. [Polly Huang](#)

Bachelor of Science in Electrical Engineering

Sep. 2015 – Jan. 2020

RESEARCH EXPERIENCE

Internet Initiative Japan

Research Intern, Reported to Dr. [Romain Fontugne](#) & Dr. [Kenjiro Cho](#)

Tokyo, Japan

Aug. 2024 – Sep. 2024 (expected)

* **Mapping Critical Network Infrastructure:** The Internet is built upon layers of critical network infrastructure, including Internet Exchange Points (IXPs) and submarine cables. In this analysis, we examine how different regions across the world depend on IXPs and submarine cables to access popular content and Internet services. We also highlight the risks associated with network routes concentrating around a few critical pieces of infrastructure.

Cisco ThousandEyes

Research Intern, Reported to Dr. [Kyle Schomp](#) & Dr. [Arash Molavi](#)

San Francisco, CA, USA

May. 2024 – Aug. 2024

* **Bringing State into IP geolocation:** IP geolocation is the process of mapping a logical IP address to a geographical entity. Latency-based approaches typically rely on stateless evaluations of Round-Trip-Time (RTT) from known vantage points. In this study, we investigate the advantages of integrating state information into the geolocation process and assess the effects of reducing the confidence of vantage point locations.

AquaLab, Northwestern University

Research Assistant, Advised by: Prof. [Fabian E. Bustamante](#)

Evanston, IL, USA

Sep. 2022 – Present

* **Submarine Cable Network Criticality:** The submarine cable network is a critical infrastructure that underpins the global Internet. However, little is known about the importance of specific cables and landing stations within a given region. In this project, we leveraged submarine and terrestrial cable maps to develop a system that maps network-layer measurements to their underlying network infrastructure and derived metrics to evaluate the relative importance of submarine cables and landing stations in different regions.

* **Investigation of IP geolocation Techniques:** In an attempt to develop a local IP geolocation service, we implemented an active geolocation technique (`single-radius`), and evaluated the efficacy of it's vantage point selection strategy. We integrated efforts from `Manycast2`, `sc.hoiho`, and `single-radius` to build a geolocation pipeline that focuses on core Internet infrastructure.

Network & Systems Lab, National Taiwan University

Research Assistant, Advised by Prof. [Polly Huang](#)

Taipei, Taiwan

Jan. 2020 – Jun. 2021

* **Exploring Twitch's Video Delivery Infrastructure:** Twitch is a global live streaming platform primarily focused on video gaming. In this project we built `Jujuby` to measure Twitch's video delivery mechanism at scale. We deployed Jujuby on over 20 vantage points across the globe and performed a longitudinal analysis of the scale and coverage of Twitch's Content Delivery Network (CDN).

Speech Processing Lab, National Taiwan University

Undergraduate Researcher, Advised by Prof. [Lin-Shan Lee](#)

Taipei, Taiwan

Sep. 2017 – Jul. 2018

* **Speech Transformation:** Developed a generative Variational Autoencoder (VAE) model using a sequence-to-sequence architecture with GRUs and attention layers to transform the speaker identity in uttered speech, and conducted an analysis on the interpretability of latent variables within VAEs to enhance prosody conversion

PUBLICATIONS

- [C1] Ten years of the Venezuelan crisis - An Internet perspective. *Esteban Carisimo, Rashna Kumar, **Caleb J. Wang**, Santiago Klein, Fabian E. Bustamante*. In *proceedings of the ACM SIGCOMM 2024 Conference, Sydney, Australia*
- [C2] A Hop Away from Everywhere: A View of the Intercontinental Long-haul Infrastructure. *Esteban Carisimo, **Caleb J. Wang**, Mia Weaver, Fabian E. Bustamante, Paul Barford*. In *the proceedings of the ACM on Measurement and Analysis of Computing Systems (SIGMETRICS'24), Venice, Italy*
- [C3] Jujuby: Design and Deployment of a Crawler for Twitch CDN Mapping. ***Caleb Wang**, Yuan-Tai Lui, Polly Huang*. In *proceedings of the 17th Asian Internet Engineering Conference (AINTEC'22), Hiroshima, Japan*
- [P1] Towards Cost Effective Server Population Estimation: A Case Study of Twitch. ***Caleb Wang**, Hsi Chen, Polly Huang*. In *the proceedings of the Web Conference 2020 (WWW'20), Taiwan, Taipei*
- [P2] Twitch's CDN as an Open Population Ecosystem. *Wei-Shiang Wung, **Caleb Wang**, Hsi Chen, Yuan-Tai Lui, Cheng Hsu, Guan-Ting Ting, Polly Huang*. 2020. *ACM SIGCOMM 2020 Conference*

WORK EXPERIENCE

Microsoft Taiwan

Customer Engineer

Taipei, Taiwan

Aug. 2021 – Jun. 2022

* **GCR Function - Customer Success:** As an engineer under the Customer Success Unit, I was responsible for proactively engaging with customers to ensure the effective adoption of Microsoft services and solutions. I provided technical consultation on SQL Server and PowerBI for Taiwanese SMC and government agencies and developed custom IPs for Microsoft Synapse and Databricks.

AIR&D Center, Microsoft Taiwan

Software Engineer Intern

Taipei, Taiwan

Sep. 2019 – Jun. 2021

* **Document Intelligence:** As a software engineer intern, I implemented a CI pipeline to ensure synchronization of containerized environments across development teams and reduced 50% testing overhead by automating the testing CI/CD pipeline for production containers.

* **Partner Technology, Hong Kong:** As part of the Partner Technology team, I worked with Microsoft partners in Hong Kong on developing custom AI solutions, leading to case wins in financial, insurance, legal, and government sectors. I reduced service evaluation overhead from days to minutes by implementing a demonstration platform for partners to quickly evaluate Azure services.

Yoctol.AI

Data Scientist Intern

Taipei, Taiwan

Jul. 2018 – Aug. 2018

* **Speaker Identification:** As a data scientist intern, I trained a speaker identification model based on a Siamese architecture with Tensorflow, and hosted the model as a backend for a web-based speaker identification service using Flask, Redis and Tensorflow serving.

TEACHING EXPERIENCE

CS 345: Distributed Systems, Northwestern University

Teaching Assistant

Spring 2024

Designed lecture materials, graded programming assignments, and held weekly office hours to assist students.

Computer Networks, National Taiwan University

Teaching Assistant

Fall 2020

Designed lecture materials, graded programming assignments, and held weekly office hours to assist students.

Machine Learning, National Taiwan University

Head Teaching Assistant

Fall 2018 & Fall 2019

As head teaching assistant, I built an online grading system to automatically grade assignments and rank scores. The system consists of an inference engine (*Flask, PyTorch*), OAuth login service (*Node.js*), a database (*MongoDB*), and a web server (*Nginx*), and was horizontally scaled with *Docker, docker-compose* to support more than 200 students.

SKILLS

Languages: Python, SQL, JavaScript, C/C++, Bash

Frameworks & Libraries: React, Node.js, Flask, Django, PyTorch, MongoDB, Redis

DevOps: Azure DevOps, Git, Docker