









Scott Freitas


I'm a Principal Applied Scientist at Microsoft working at the intersection of applied and theoretical machine learning, with a focus on **graph mining** and **deep learning**. My goal is to develop explainable, robust, and efficient next-generation cybersecurity systems.

I completed my Machine Learning PhD at [Georgia Tech](#) where I worked with [Polo Chau](#). I co-authored several winning research proposals, including a multi-million dollar [DARPA grant](#); was awarded PhD fellowships from [IBM Research](#), [NSF GRFP](#) and [Raytheon](#); and was fortunate to work with amazing researchers at [IBM Research](#), [Amazon](#), [Microsoft Advanced Threat Protection](#), [Microsoft Research](#), [Intel](#) and the [Naval Air Warfare Center](#).

 scottfreitas.com
 safreita1@gmail.com
 [Google Scholar](#)
 [Blog Posts](#)
 [Curriculum Vitae \(PDF\)](#)

 [Github](#)
 [Linkedin](#)
 [@scottafreitas](#)
 [YouTube](#)

Education

- Dec. 2021 **Ph.D. in Machine Learning**
Aug. 2018 Georgia Institute of Technology, Atlanta, GA
Advisor: Duen Horng (Polo) Chau
Thesis: *Developing Robust Models, Algorithms, Databases and Tools with Applications to Cybersecurity and Healthcare*
Committee: Duen Horng (Polo) Chau, Srijan Kumar, Diyi Yang, B. Aditya Prakash, Hanghang Tong
 [Thesis](#)  [Thesis Recording \(Proposal\)](#)  [Thesis Slides](#)
- May 2018 — **M.S. in Computer Science**
May 2017 Arizona State University, Tempe, AZ
Advisor: Hanghang Tong
Thesis: *Mining Marked Nodes in Large Graphs*
Committee: Hanghang Tong, Ross Maciejewski, Yezhou Yang
GPA: 4.00/4.00
 [Thesis](#)
- May 2017 — **B.S. in Computer Science**
Aug. 2015 Arizona State University, Tempe, AZ
Advisor: Ross Maciejewski
Thesis: *Guided Augmented Reality Tours using Landmarks and Social Media*
GPA: 3.98/4.00
 [Thesis](#)  [Thesis Recording](#)

May 2014 — **B.S.E. in Electrical Engineering**
Aug. 2010 Arizona State University, Tempe, AZ
Advisor: James Aberle
Thesis: *Multi-Stage Linear Electromagnetic Accelerator Using Optical Triggering*
GPA: 3.64/4.00
[📄 Thesis](#) [🎧 Thesis Recording](#)

Honors and Awards

2021 **IBM PhD Fellowship**
One of sixteen fellows; awarded for my work in developing next-generation explainable defenses

2021 **Nvidia Data Science Teaching Kit**
Helped develop one of five Nvidia teaching kits used by educators around the world

2019 **Raytheon Research Fellowship**
Awarded for my PhD work in adversarial machine learning

2018 — 2021 **NSF Graduate Research Fellowship**
National Science Foundation recognizes and supports outstanding graduate students in STEM fields

2018 **Outstanding Computer Science Masters Student (ASU)**
Awarded to single master student demonstrating exemplary performance

2017 **Best Demo Award, Runner Up at CIKM '17**
For "Rapid Analysis of Network Connectivity"

2017 **CIKM Travel Grant**
Funding from NSF and SIGWEB to present at CIKM

2016 — 2017 **FURI Grant**
Undergraduate research grant awarded for work in network connectivity

2016 — 2017 **Arizona Graduate Scholar Award**
Merit scholarship awarded to select number of master students

2010 — 2014 **Provost's Scholarship**
Merit scholarship awarded to select number of incoming undergraduate students

Industry Research Experience

Present — **Microsoft**, Redmond, WA
Sep. 2024 *Principal Applied Scientist (level 65), Microsoft Security Research*

- Leading research into LLM-based agents to automatically identify detection and disruption rule gaps.
- Designed and deployed TITAN, a graph-based threat intelligence framework integrated into Microsoft Defender XDR, achieving a 21% increase in incident disruption, a 1.9x reduction in response time, and 99% precision.
[📄 Paper](#) [📄 Blog](#) [▶ Microsoft Ignite Talk](#)
- Developed an adaptive incident prioritization score that assists analysts in prioritizing security incidents for investigation.

Aug. 2024 **Microsoft**, Redmond, WA
Sep. 2023 *Senior Applied Scientist (level 64), Microsoft Security Research*

- Led an ML research team in architecting and delivering key capabilities for our flagship AI product, Copilot for Security, including tailored recommendations for similar incidents, triaging, and remediation. Collaborated across teams to launch the product on a tight timeline.
[📄 Paper](#) [📄 Blog](#) [📄 Dataset](#)

- Developed GraphWeaver, a geo-distributed alert correlation framework integrated into Microsoft Defender XDR, achieving 99% correlation accuracy while handling billions of security alerts across hundreds of thousands of enterprises. Reduced our singleton incident rate by 7%, translating into millions of investigation hours saved annually by SOCs.

[Paper](#) [Blog](#)

Aug. 2023

Microsoft, Redmond, WA

Jan. 2022

Senior Applied Scientist (level 63), Microsoft Security Research

- Developed graph-based algorithms to identify alert correlation gaps, enabling the correlation of millions of alerts into comprehensive incident stories, saving customers millions in investigation time.
- Led the development and execution of a comprehensive research integration plan, successfully help merge two billion-dollar security products, M365D and Sentinel, into Microsoft Defender XDR.

[Blog](#)

Dec. 2021 —

IBM Research, Yorktown Heights, NY

Sep. 2021

Research Intern, Cyber Security Intelligence (CSI) Team

Mentor: Teryl Taylor, Frederico Araujo, Jiyong Jang

Developed unsupervised graph representation learning techniques to detect suspicious activity in cloud platforms

Aug. 2021 —

Amazon, Seattle, WA

May 2021

Applied Scientist Intern, Fraud Detection and Risk Transaction (CTPS)

Mentor: Hao Zheng, Yanni Lai

Created unsupervised and semi-supervised approaches to prevent fraudulent transactions across the Amazon marketplace

May 2020 —

Microsoft, Redmond, WA

Aug. 2020

Research Intern, Microsoft ATP + Microsoft Research

Mentor: Karishma Sanghvi, Yuxiao Dong

Designed semi-supervised graph neural network approach to detect malicious software

Aug. 2019 —

Microsoft, Redmond, WA

May 2019

Research Intern, Microsoft Advanced Threat Protection (ATP)

Mentor: Andrew Wicker, Joshua Neil

• Created first framework to model lateral attacks on enterprise networks, enabling IT admins to quantify and mitigate network vulnerability to lateral attacks

[Paper](#)

March 2015 —

General Dynamics, Scottsdale, AZ

Dec. 2014

Systems Engineer, Mission Systems

Worked on the Integrated Threat Force team to develop and refine the communication technology systems.

Aug. 2013 —

Naval Air Warfare Center, Point Mugu, CA

May 2013

Research Intern, Naval Research Enterprise Internship Program (NREIP)

Mentor: Balaji Iyer

Explored methods of preventing electromagnetic interference from coupling into superconducting receivers

Academic Research Experience

Present —

Georgia Institute of Technology, Atlanta, GA

Aug. 2018

Graduate Research Assistant, School of Computational Science and Engineering

Mentor: Duen Horng (Polo) Chau

Member of the Polo Club of Data Science where we innovate scalable, interactive, and interpretable tools that amplify human's ability to understand and interact with billion-scale data and machine learning models

May 2018 — **Arizona State University**, Tempe, AZ

Summer 2017 *Graduate Research Assistant, School of Computing, Informatics, and Decision Systems Engineering*

Mentor: Hanghang Tong

Conducted research in graph based connectivity analysis to improve local graph partitioning. Developed web-based prototype for explainable ranking in complex multi-layered networks.

Aug. 2017 — **Arizona State University**, Tempe, AZ

May 2017 *Summer Research Assistant, School of Computing, Informatics, and Decision Systems Engineering*

Mentor: Ross Maciejewski

Developed interactive augmented reality (AR) graph models in the Microsoft HoloLens.

May 2017 — **Arizona State University**, Tempe, AZ

Jan. 2016 *Undergraduate Research Assistant, School of Computing, Informatics, and Decision Systems Engineering*

Mentor: Hanghang Tong

Developed fast graph mining algorithms for network connectivity analysis, and award winning web platform for visualization and analysis.

Publications

Web Scale Graph Mining for Cyber Threat Intelligence

Scott Freitas, Amir Gharib

arXiv (arXiv). 2024.

[Project](#) [PDF](#) [Video](#) [BibTeX](#) [🏆 Deployed in Microsoft Unified Security Operations Platform 🏆](#)

[Presented at Microsoft Ignite 2024](#)

AI-Driven Guided Response for Security Operation Centers with Microsoft Copilot for Security

Scott Freitas, Jovan Kalajdjieski, Amir Gharib, Rob McCann

arXiv (arXiv). 2024.

[Project](#) [PDF](#) [Blog](#) [Dataset](#) [BibTeX](#) [🏆 Deployed in Microsoft Copilot for Security product](#)

GraphWeaver: Billion-Scale Cybersecurity Incident Correlation

Scott Freitas, Amir Gharib

ACM International Conference on Information and Knowledge Management (CIKM). Boise, Idaho, 2024.

[Project](#) [PDF](#) [Blog](#) [BibTeX](#) [🏆 Deployed in Microsoft Defender XDR product 🏆 Keynote talk at CIKM](#)

[Industry Day](#)

Graph Vulnerability and Robustness: A Survey

Scott Freitas, Diyi Yang, Srijan Kumar, Hanghang Tong, Duen Horng (Polo) Chau

IEEE Transactions on Knowledge and Data Engineering (TKDE). 2022.

[PDF](#) [BibTeX](#)

MalNet: A Large-Scale Image Database of Malicious Software

Scott Freitas, Rahul Duggal, Duen Horng (Polo) Chau

ACM International Conference on Information and Knowledge Management (CIKM). Atlanta, GA, 2022.

[Demo](#) [PDF](#) [Dataset](#) [Code](#) [BibTeX](#)

A Large-Scale Database for Graph Representation Learning

Scott Freitas, Yuxiao Dong, Joshua Neil, Duen Horng (Polo) Chau

Neural Information Processing Systems Datasets and Benchmarks (NeurIPS). Virtual, 2021.

[Project](#) [Demo](#) [PDF](#) [Blog](#) [Dataset](#) [Code](#) [BibTeX](#)

Evaluating Graph Vulnerability and Robustness using TIGER

Scott Freitas, Diyi Yang, Srijan Kumar, Hanghang Tong, Duen Horng (Polo) Chau

ACM International Conference on Information and Knowledge Management (CIKM). Virtual, 2021.

[PDF](#) [Blog](#) [Video](#) [Code](#) [BibTeX](#) [Featured in Nvidia Data Science Toolkit](#)

EnergyVis: Interactively Tracking and Exploring Energy Consumption for ML Models

Omar Shaikh, Jon Saad-Falcon, Austin P Wright, Nilaksh Das, Scott Freitas, Omar Asensio, Duen Horng Chau

ACM Conference on Human Factors in Computing Systems (CHI). Virtual, 2021.

[Demo](#) [PDF](#) [Video](#) [Code](#) [BibTeX](#)

UnMask: Adversarial Detection and Defense Through Robust Feature Alignment

Scott Freitas, Shang-Tse Chen, Zijie J. Wang, Duen Horng (Polo) Chau

IEEE International Conference on Big Data (Big Data). Atlanta, GA, 2020.

[Project](#) [PDF](#) [Blog](#) [Video](#) [Code](#) [BibTeX](#)

HAR: Hardness Aware Reweighting for Imbalanced Datasets

Rahul Duggal, Scott Freitas, Sunny Dhamnani, Duen Horng (Polo) Chau, Jimeng Sun

IEEE Conference on Big Data (Big Data). Orlando, USA, 2021.

[PDF](#) [Video](#) [BibTeX](#)

Argo Lite: Open-Source Interactive Graph Exploration and Visualization in Browsers

Siwei Li, Zhiyan Zhou, Anish Upadhayay, Omar Shaikh, Scott Freitas, Haekyu Park, Zijie J. Wang, Susanta Routray, Matthew Hull, Duen Horng (Polo) Chau

ACM International Conference on Information and Knowledge Management (CIKM). Virtual, 2020.

[Demo](#) [PDF](#) [Code](#) [BibTeX](#)

REST: Robust and Efficient Neural Networks for Sleep Monitoring in the Wild

Rahul Duggal*, Scott Freitas*, Cao Xiao, Duen Horng (Polo) Chau, Jimeng Sun

The Web Conference (WWW). Taipei, Taiwan, 2020.

[Project](#) [PDF](#) [Blog](#) [Video](#) [Code](#) [BibTeX](#) * Authors contributed equally

D²M: Dynamic Defense and Modeling of Adversarial Movement in Networks

Scott Freitas, Andrew Wicker, Duen Horng (Polo) Chau, Joshua Neil

SIAM International Conference on Data Mining (SDM). Cincinnati, Ohio, 2020.

[Project](#) [PDF](#) [Blog](#) [BibTeX](#)

Extracting Knowledge For Adversarial Detection and Defense in Deep Learning

Scott Freitas, Shang-Tse Chen, Duen Horng (Polo) Chau

KDD Workshop: Learning and Mining for Cybersecurity (LEMINGS). Anchorage, Alaska, 2019.

[PDF](#) [BibTeX](#)

Local Partition in Rich Graphs

Scott Freitas, Nan Cao, Yinglong Xia, Duen Horng (Polo) Chau, Hanghang Tong

IEEE International Conference on Big Data (Big Data). Seattle, Washington, 2018.

[Project](#) [PDF](#) [BibTeX](#)

X-Rank: Explainable Ranking in Complex Multi-Layered Networks

Jian Kang*, Scott Freitas*, Haichao Yu, Yinglong Xia, Hanghang Tong

ACM International Conference on Information and Knowledge Management (CIKM). Turin, Italy, 2018.

[Project](#) [PDF](#) [BibTeX](#) * Authors contributed equally

Rapid Analysis of Network Connectivity

Scott Freitas, Hanghang Tong, Nan Cao, Yinglong Xia

ACM International Conference on Information and Knowledge Management (CIKM). Singapore, 2017.

[Project](#) [PDF](#) [Video](#) [Code](#) [BibTeX](#) [Best Demo Paper, Runner up](#)

Datasets and Tools

2024 **GUIDE:** Largest public collection of real-world cybersecurity incidents

Scott Freitas, Jovan Kalajdjieski, Amir Gharib, Rob McCann

[Dataset](#)

2022 **MalNet-Image:** Largest dataset for image-based malware classification

Scott Freitas, Rahul Duggal, Duen Horng (Polo) Chau

[Dataset](#)

2021 **MalNet-Graph:** Largest dataset for graph representation learning and classification

Scott Freitas, Yuxiao Dong, Joshua Neil, Duen Horng (Polo) Chau

[Dataset](#)

2021 **TIGER:** Comprehensive Python toolbox to evaluate graph vulnerability and robustness

Scott Freitas, Diyi Yang, Srijan Kumar, Hanghang Tong, Duen Horng (Polo) Chau

[Code](#)

Patents

2024 **Threat Actor Infrastructure Profiling Using a Graph and Reputation Propagation** (Filed)

Scott Freitas, Amir Gharib

Microsoft

2024 **Adaptive Incident Prioritization Engine in a Security Management System** (Filed)

Scott Freitas, Amir Gharib

Microsoft

2024 **Geographically Diversified Embedding-Based Guided Response to a Security Alert** (Filed)

Scott Freitas, Jovan Kalajdjieski, Amir Gharib, Rob McCann

Microsoft

2024 **Cybersecurity Incident Correlation** (Filed)

Scott Freitas, Amir Gharib

Microsoft

2023 **Hierarchical Representation Models** (Filed)

Jovan Kalajdjieski, Scott Freitas, Amir Gharib, Rob McCann

Microsoft

Talks

Simplify your SOC with Rob Lefferts and Allie Mellen

Nov. 2024
Microsoft Ignite

GraphWeaver: Billion-Scale Cybersecurity Incident Correlation

Oct. 2024
Research Paper Invited for Keynote Talk at CIKM Industry Day

Clustering Process Activity in Cloud Environments using Graph Representation Learning

Dec. 2021
IBM Research

Dec. 2021
DARPA CHASE: Cyber-Hunting at Scale

Detecting Financial Fraud in Online Marketplaces

August 2021
Amazon

Developing Robust Models, Algorithms, Databases and Tools with Applications to Cybersecurity and Healthcare

October 2021
GE Research

Dec. 2021
Georgia Institute of Technology

May 2021
Georgia Institute of Technology

Exploring Graph Neural Networks for Malware Detection

July 2020
Microsoft Advanced Threat Protection

On the Robustness and Vulnerability of Graphs

April 2020
Georgia Institute of Technology

D²M: Dynamic Defense and Modeling of Adversarial Movement in Networks

Aug. 2019
Microsoft Advanced Threat Protection Research Expo

Mining Marked Nodes in Large Graphs

Dec. 2018
Microsoft Advanced Threat Protection Group

May 2018
Arizona State University

Local Partition in Rich Graphs

Dec. 2018
IEEE International Conference on Big Data

Rapid Analysis of Network Connectivity

Nov. 2017
ACM International Conference on Information and Knowledge Management (CIKM)

Network Connectivity Analysis and Visualization in Large Graphs

April 2017
Keynote Speaker: ASU Fulton Undergraduate Research Initiative (FURI)

Nov. 2016
ASU FURI Research Symposium

Press

Nov. 2024
"Ignite news: What's new in Microsoft Defender XDR?",

Sept. 2024
"AI-Driven Guided Response for SOCs with Microsoft Copilot for Security",

August 2024
"Cybersecurity incident correlation in the unified security operations platform",

April 2024
"Triage and investigate incidents with guided responses from Microsoft Copilot in Microsoft Defender",

Dec. 2021	"Congratulations to the Newest PhDs from Georgia Tech",
June 2021	"New NVIDIA Partnership Bridges Education Gap for Data Science and Machine Learning",
April 2021	"ML Student Earns Prestigious IBM Ph.D. Fellowship Award",
April 2021	"IBM PhD Fellowship Awardees Announced",
April 2021	"Accelerated Data Science in the Classroom: Teaching Analytics and Machine Learning with RAPIDS",
April 2020	"Georgia Tech and Intel Awarded Multimillion-Dollar Program to Defend Against Attacks on AI",
April 2020	"DARPA Snags Intel to Lead its Machine Learning Security Tech",
April 2020	"Machine Learning Technique Helps Wearable Devices Get Better at Diagnosing Sleep Disorders and Quality",
Feb. 2019	"Raytheon Awards Two ML@GT Students Graduate Research Assistantships",
July 2018	"NSF Graduate Research Fellow wants to use computer science to solve society's toughest problems",

Grants and Funding

2021	<p>IBM PhD Fellowship IBM Research PhD Fellowship Awardee Funded: \$95,000</p>
2020	<p>Google Cloud Research Grant Large Scale Malware Analysis Funded: \$5,000 Google cloud credits</p>
2018 — 2022	<p>Guaranteeing AI Robustness against Deception (GARD) DARPA Research Grant Co-PIs: Jason Martin, Duen Horng (Polo) Chau Funded: multi-million Helped formulate adversarial defense techniques</p>
2018	<p>Amazon AWS Research Grant Adversarial Re-Training and Model Vaccination for Robust Deep Learning Funded: \$5,000 AWS cloud credits</p>
2018	<p>Nvidia GPU Grant Defending Adversarial Attacks by Robust, Inference-time Local Linear Approximation Funded: Nvidia Titan V GPU worth \$3,000</p>
2019	<p>Raytheon Research Fellowship Extracting Knowledge For Adversarial Detection and Defense Funded: \$25,000</p>
2018 — 2023	<p>NSF Graduate Research Fellowship Program (GRFP) Multi-level Interdiction and Assistance Modeling for Natural Disasters Funded: Full tuition + \$102,000</p>
2016 — 2017	<p>FURI Grant Network Connectivity Analysis and Visualization in Large Graphs Funded: \$3,000</p>

Teaching

- Spring 2021 **Graduate Teaching Assistant**
Georgia Institute of Technology, Atlanta, GA
Data and Visual Analytics, Instructor: Duen Horng (Polo) Chau
- Fall 2020 **Graduate Teaching Assistant**
Georgia Institute of Technology, Atlanta, GA
Data and Visual Analytics, Instructor: Duen Horng (Polo) Chau
- Fall 2013 **Undergraduate Teaching Assistant**
Arizona State University, Tempe, AZ
Fulton Undergraduate Research Experience (FSE 294), Instructor: Joshua Lyon
Designed and taught introductory lesson plans to new engineering students

Mentoring

- Summer 2023 **Davinder Kaur** at Microsoft
Ph.D. in Computer Science, Indiana University–Purdue University Indianapolis
- Summer 2023 **Joshua Feinglass** at Microsoft
Ph.D. in Computer Engineering, Arizona State University
- Fall 2020 **Kevin Li**
Summer 2020 *B.S. in Computer Science, Georgia Institute of Technology*
- Fall 2020 **Omar Shaikh**
Spring 2020 *B.S. in Computer Science, Georgia Institute of Technology*
- Fall 2020 **Jon Saad-Falcon**
Spring 2020 *B.S. in Computer Science, Georgia Institute of Technology*
- Fall 2020 **Frank Zhou**
Spring 2020 *B.S. in Computer Science, Georgia Institute of Technology*

Service

Hiring Committee

- Microsoft Security Research (**Microsoft**) 2024
Microsoft Security Research Summer Interns (**Microsoft**) 2022-2024

Program Committee

- Association for the Advancement of Artificial Intelligence (**AAAI**) at AAAI 2021
ACM International Conference on Information and Knowledge Management (**CIKM**) at ACM CIKM 2020

Reviewer

- Practice of Knowledge Discovery in Databases (**ECML-PKDD**) 2021
International Conference on Computer Vision (**ICCV**) 2021
Conference on Computer Vision and Pattern Recognition (**CVPR**) 2021

ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD**) 2019, 2025

International Conference on Machine Learning (**ICML**) 2019

Technology Skills

OS and Tools: Ubuntu, Unix command line, Windows, PyCharm, Azure, Synapse, Git, Latex, AWS EC2

Programming: Python, PySpark, Kusto, SQL, Matlab, Java, C#, C++, JavaScript, HTML

Research: Machine learning, Data mining, Graph mining, Data science, Artificial intelligence, Generative AI, Large language models (LLMs), Deep learning, Computer vision, Natural language processing (NLP), Anomaly detection, Cybersecurity