Yasanka Sameera Horawalavithana

5022, Sunridge Palms Dr. Apt 204, Tampa, FL 33617 US
Email: sameera1@usf.edu
Web: https://samtube405.github.io/_profile
TP: +1 (813) 409-9303

OBJECTIVE

My broad research interests are in Foundation Models (Language + Multimodal), Geometric Deep Learning (Dynamic Graphs), Privacy and Security. To this end, I have developed multiple analytical tools to understand cyber-social systems from the quantitative and/or computational perspective.

EDUCATION

University of South Florida, Tampa, FL, USA
Ph.D., Computer Science and Engineering, August 2021 GPA: 3.96

University of Colombo, School of Computing, Colombo, Sri Lanka
Bachelor of Science (Hons.), Computer Science, April 2015, (Top 1%) GPA: 3.84

Umeå University, Sweden
Exchange Student, Computer Science and Engineering, Fall 2013

EXPERIENCE

Research Scientist Pacific Northwest National Laboratory
Richland, WA May 2022 - Present
Working at AI Group, Physical and Computational Science Directorate

Postdoctoral Research Associate Pacific Northwest National Laboratory
Richland, WA September 2021 - April 2022
Worked at National Security Directorate under Svitlana Volkova

Graduate Research Assistant University of South Florida
Department: CSE August 2016 - August 2021
Worked at Distributed System Group (DSG) under Adriana Iamnitchi.

Software Engineer Sysco Labs (Pvt.) Ltd.
59, Flower Rd, Colombo 07, Sri Lanka February 2015 - July 2016
Worked at Data and Analytic Team.

PNNL PROJECT HIGHLIGHTS

Mega-AI: Foundation Models for Science and Security (LDRD)[1] Mega AI will develop foundational AI models across multiple modalities to augment scientists for solving science problems. My contribution is for training large-scale language models from millions of scholarly documents: papers, patents, reports, books, etc.. I am leading a team of experts to show the usability of AI models for science problems focusing on molecular chemistry. *Technologies: Pytorch, Natural Language Processing, Generative Pre-trained Transformer (GPT), Distributed Training Frameworks (DeepSpeed, Megatron)* (October 2021 - Present)

EXPERT: Global-Scale Cross-Lingual Proliferation Expertise Identification and Global Expertise Forecasting We aim to forecast the global expertise and capability evolution in various scientific communities. My contribution is for developing a new deep learning architecture using graph neural network (*Dynamic Graph Transformers*) to model dynamic heteregenous graphs constructed from scientific article texts. *Technologies: Pytorch, Pytorch Geometric, Graph Neural Networks with Transformers, Knowledge Graph Forecasting* (September 2021 - Present)

PANDA: AI-Driven Predictive Analytics to Enhance Nuclear Proliferation Detection in Urban Environments PANDA will develop novel AI-driven predictive analytical framework to anticipate background isotope signatures across locations and detectors to mitigate nuance alarms. My contribution is for pretraining a graph-based machine learning model that encodes dynamic signals captured from urban sensing

[1] https://www.pnnl.gov/projects/mega-ai

systems for forecasting radiological isotopes signatures. *Technologies: Pytorch, Self-supervised Learning with Graph Neural Networks* (December 2021 - Present)

PROJECTS & PUBLICATIONS

Modeling Information Diffusion Processes with Deep Learning Algorithms (SocialSim, DARPA)[2], The objective of this work is to develop technologies for high-fidelity simulation of online social behavior (the spread and evolution of online information) while rigorously testing and measuring simulation accuracy in operational scenarios defined by Department of Defense, USA. We are designing novel methods to model the structure of information cascades in large scale social networks using Deep Learning techniques. *Technologies: Python, Tensorflow, Keras, Networkx, iGraph, Time-series Forecasting, Information Cascade Prediction, Influence Operations, Dynamic Graph Generation, Large-scale Social Network Data Analysis (Twitter, YouTube, Reddit, GitHub, Telegram)* (Jan 2018 - August 2021)

→ **Data-driven Simulation**

1. Sameera Horawalavithana Data-driven Studies on Social Networks: Privacy and Simulation, Ph.D. Dissertation, Department of Computer Science and Engineering, University of South Florida, June 2021 (**Nomination for the Outstanding Thesis and Dissertation Award**)
2. Sameera Horawalavithana, Nazim Choudhury, and Adriana Iamnitchi, Online Discussion Threads as Conversation Pools: Predicting the Growth of Discussion Threads on Reddit, Computational & Mathematical Organization Theory (CMOT).
3. Kin Wai NG, Sameera Horawalavithana, and Adriana Iamnitchi, Social-Media Activity Forecasting with Exogenous Information Signals, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), The Hague, Netherlands, 2021.
4. Renhao Liu, Fredrick Mubang, Lawrence O. Hall, Sameera Horawalavithana, Adriana Iamnitchi and John Skvoretz, Predicting Longitudinal User Activity at Fine Time Granularity in Online Collaborative Platforms, 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, 2019.

→ **Graph Analytics**

1. Sameera Horawalavithana, Kin Wai NG, and Adriana Iamnitchi, Drivers of Polarized Discussions on Twitter during Venezuela Political Crisis, The 13th International ACM Conference on Web Science (WebSci), 2021.
2. Sameera Horawalavithana, Ravindu De Silva, Mohamed Nabeel, Charitha Elvitigala, Primal Wijesekara, and Adriana Iamnitchi, Malicious and Low Credibility URLs on Twitter during the AstraZeneca COVID-19 Vaccine Development, International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS), DC, USA, 2021 **(Best COVID Track Paper Award)**
3. Kin Wai NG, Sameera Horawalavithana, and Adriana Iamnitchi, Multi-platform Information Operations: Twitter, Facebook and YouTube against the White Helmets, The Workshop Proceedings of the 14th International AAAI Conference on Web and Social Media, 2021.
4. Sameera Horawalavithana, Kin Wai NG, and Adriana Iamnitchi,. Twitter is the Megaphone of Cross-Platform Messaging on the White Helmets, International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS), DC, USA, 2020
5. Sameera Horawalavithana, Abhishek Bhattacharjee, Renhao Liu, Nazim Choudhury, Lawrence O. Hall, and Adriana Iamnitchi,. Mentions of Security Vulnerabil-

[2] http://www.cse.usf.edu/socialsim/

ities in Reddit, Twitter and GitHub, IEEE/WIC/ACM International Conference on Web Intelligence, Thessaloniki, Greece, October, 2019

Structural Anonymization Techniques for Large, Labeled, and Dynamic Social Graphs (NSF), The objective of this work is to provide big data owners with tools to safely share their social networks data with the research community. The project aims to approach graph anonymization via two techniques for graph generation: dK-series techniques, and Exponential Random Graph Model-based approaches (ERGM). *Technologies: Python, R, Random Graph Generative Models, Statistical Models, Causality Analysis* (Aug 2016 - Aug 2020)

1. Sameera Horawalavithana, Juan Arroyo Flores, John Skvoretz, and Adriana Iamnitchi. Behind the Mask: Understanding the Structural Forces that Make Social Graphs Vulnerable to De-anonymization. IEEE Transactions on Computational Social Systems (TCSS), 2019
2. Sameera Horawalavithana, Juan Arroyo Flores, John Skvoretz, and Adriana Iamnitchi. The Risk of Node Re-identification in Labeled Social Graphs, Applied Network Science (2019)
3. Sameera Horawalavithana, Clayton Gandy, Juan Arroyo Flores, John Skvoretz, and Adriana Iamnitchi. Diversity, Topology, and the Risk of Node Re-identification in Labeled Social Graphs., The 7^{th} International Conference on Complex Networks and Their Applications. Cambridge, UK, Dec. 2018

Group Dynamics in Online Games Modeling the behavior of online players to understand how groups form, and evolve dynamically (December 2016 - August 2019)

1. Essa Alhazmi, Sameera Horawalavithana, Jeremy Blackburn, John Skvoretz and Adriana Iamnitchi. An Empirical Study on Team Formation in Online Games. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Australia, 2017
2. Essa Alhazmi, Naim Choudhury, Sameera Horawalavithana and Adriana Iamnitchi. Temporal Mobility Networks in Online Gaming. Frontiers in Big Data, 2019.

Top-k publish/subscribe model: Developed a content-based Top-k publish/subscribe model with a novel formalization of a ranking method that uses a variation of minimum-independent dominating set problem in dynamic graphs. Randomized algorithm is proposed to rank streaming contents. *Technologies: Java, Amazon Kinesis, Amazon ElasticCache.* (March 2014 - December 2014)

1. Sameera Horawalavithana and D. N. Ranasinghe. 2015. An Efficient Incremental Indexing Mechanism for Extracting Top-k Representative Queries Over Continuous Data-streams. In Proceedings of the 14th International Workshop on Adaptive and Reflective Middleware (ARM) colocated with USENIX Middleware, Vancouver, Canada, December, 2015
2. Sameera Horawalavithana, D.N. Ranasinghe. Cloud based pub/sub model for Top-k matching over continuous data-streams **(Best Undergraduate Thesis)** January 2015, University of Colombo, School of Computing, Sri Lanka

Real-time ETL: Contributed to the development of a (near) real-time data-warehouse solution during my tenure at Sysco Labs, *Technologies: Java, Apache Storm, Hadoop Eco-system, MySQL* (February 2015 - August 2015)

EXTERNAL GRANT PROPOSALS (PNNL)

- Reproducible, Modular, multidimensional, and interpretable HIATUS Evaluation, December, 2021 (PI: Dr. Svitlana Volkova). My contribution is within the privacy preservation metrics (TA3) in the statement of work (SoW) submitted for the Human Interpretable Attribution of Text using Underlying Structure (HIATUS) Testing and Evaluation (T&E) proposal call by the Office of the Director of National Intelligence, Intelligence Advanced Research Projects Activity (IARPA).

INVITED TALKS

- Lessons from Developing Data-driven Simulators of Information Spread in Social Media, Networks, 2021
- Generative/Discriminative Approach to De-construct Cascading Events, Machine Learning in Network Science, NetSci, 2019

AWARDS & HONORS

- Nomination for the Outstanding Thesis and Dissertation Award, 2021/2022, University of South Florida
- Best COVID Track Paper Award, International Conference on Social Computing, Behavioral-Cultural Modeling, & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRiMS), DC, USA, 2021
- Winner, Grand Challenge, COVID-19 AstraZeneca Vaccine Disinformation, The 3rd North American Social Networks Conference (NASN), 2021
- Best Computer Science Undergraduate Thesis Award in the year 2014, University of Colombo, Sri Lanka
- Linnaeus Palme Scholarship on the International Exchange Programme administered by the International Programme Office for Education and Training and financed by Sida, Swedish International Development Co-operation Agency.
- Scholarship to attend Summer Institute, San Diego Super Computing Center (SDSC), University of California, San Diego, Summer, 2018
- Scholarship to attend 2^{nd} International summer school for Deep Learning, University of Genoa, Genoa, Italy, Summer, 2018
- ACM/SIGHPC Travel Grant for Supercomputing conference 2016, Utah
- Achieved 3rd place at Nordic Collegiate Programming Contest 2013 (Umeå region) representing Umeå university, Sweden
- Mahapola Merit Higher Education Scholarship, Government of Sri Lanka, 2010

TEACHING & MENTORSHIP

Google Summer of Code (GSoC) Mentorship I served as an advisor to the open-source organization, Sustainable Computing Research Lab (SCoRe), where I mentored two GSoC students to implement a crowd sourced fact checking platform. The goal of this platform (FactBounty) is to promote the civic participation to minimize the spread of false news and rumors in Sri Lankan digital eco-system.

- Anmol Bansal, Implement a backend for a crowd-driven Fact-checking platform.
- Tuan Amith, A web interface for a crowd-driven Fact-checking platform.

Research Supervision (Undergraduate)

- Milindu Sanoj Kumarage (2015), An efficient query platform for streaming and dynamic natural graphs, University of Colombo, Sri Lanka
- Malith Senaweera, Ruwanmalee Dissanayake, Nuwini Chamindi, (2017), The Influence of Community Interactions on User Affinity in Social Networks: A Facebook Case Study, University of Colombo, Sri Lanka

Teaching Assistant

Network Science, (Graduate-level, Fall, 2018), Data Structures (Undergraduate-level, Fall, 2016), University of South Florida. Prepared and evaluated homework assignments, held weekly recitations and office hours

Student Volunteer

Volunteer on organization activities at ACM Symposium on Principles of Distributed Computing (PODC) conference, Washington, DC, Summer 2017

SERVICE (PC/ Reviewing)

- ACM Web Science (WebSci) Conference, 2022
- International Workshop on Social Sensing (SocialSens), ICWSM, 2021
- IEEE Transactions on Computational Social Systems (TCSS), 2020.
- IEEE Transactions on Information Forensics & Security (TFIS), 2020
- Journal of Computational Social Science, Springer, 2020
- Transactions on Parallel and Distributed Systems (TPDS), 2018.
- Online Social Networks and Media (OSNEM), Elsevier Journal, 2018
- IEEE Conference on Distributed Computing Systems (ICDCS), 2018