Enhao Liu

Department of Enterprise Data & Analytics
University of Maryland Medical System
920 Elkridge Landing Rd.
Linthicum, MD 21090
✆ *+1(614)-313-7388*
✉ *enhao.liu@umm.edu*

Education

Aug 2017–May 2021	**Doctor of Philosophy**, *The Ohio State University*, Columbus, OH, USA.
	Industrial and Systems Engineering - Operations Research; Advisor: Theodore Allen.
Aug 2015–Aug 2017	**Master of Science**, *The Ohio State University*, Columbus, OH, USA.
	Industrial and Systems Engineering - Operations Research; Advisor: Theodore Allen.
Sep 2011–June 2015	**Bachelor of Engineering**, *Jinan University*, Guangzhou, China.
	Electric Engineering and Automation. Senior thesis advised by Weihua Li.

Skills

Programming	Python, R, SQL, Matlab, GAMS
Main Packages	Python: pyomo, Gurobi API, Cplex API, scikit-learn, imbalanced-learn, xgboost, hyperopt, keras, OpenCV, skimage, numpy, scipy, pandas, matplotlib, plotly, Flask, sqlalchemy; R: caret, DMwR, MDPtoolbox, pomdp, ompr, CVXR, dLagM, rsatscan, tidyverse, leaflet, visNetwork, odbc, shiny, shinyjs.
Software	Minitab, ARENA, Simio
Others	LaTeX, Microsoft Office Suite
Coursework	Operation Research Models and Methods; Linear Programming; Non-linear Programming; Mixed Integer Programming; Convex Optimization Algorithm; Stochastic Optimization; Stochastic Process; Decomposition Techniques in Mathematical Programming; Statistical Machine Learning; Mathematical Analysis of Algorithms; Intermediate Data Analysis.
Languages	Chinese (native), English (fluent)

Awards

2019	Runner-up, Student Paper Competition, The Social Media Analytics Section of INFORMS.
2014	Second Prize of Scholarship for Excellent Students, Jinan University.
2013	Third Prize of Scholarship for Excellent Students, Jinan University.
2013	Yihai Kerry Scholarship for Innovative Undergraduates, Jinan University.
2013	Meritorious Winner, National College Mathematical Contest in Modeling.
2012	First Prize, China Undergraduate Mathematical Contest in Modeling.

Research Interests

Methodology: (1) Machine Learning: Optimal Classification Tree, Ensemble Learning, Hyper-parameters Optimization; (2) Reinforcement Learning: Partially Observable Markov Decision Process (POMDP), Monte Carlo-Bayesian Reinforcement Learning (MC-BRL), Deep Q-Learning; (3) Simulation-based Optimization; (4) Statistic Modeling: Autoregressive Distributed Lag Models, Bayesian Hierarchical Models, Control Charts, Space-Time Scan Statistic

Applications: Cost-effective policy development in influential fields including cybersecurity and healthcare; Machine learning applications in cybersecurity and healthcare; Anomaly detection in multi-stream data; Space-time COVID-19 surveillance & interventions with simulation modeling.

I am especially interested in a challenging position that uses my machine learning, operations research, statistics, data integration, and programming skills.

Employment

May 2021–present **Healthcare Data Scientist**, *University of Maryland Medical System*, Linthicum, MD, USA.
Department: Enterprise Data & Analytics; Supervisor: D'Souza Warren

- Worked as part of a team of scientists and engineers that develop and manage predictive (machine learning and deep learning) and prescriptive (mathematical optimization and simulation) analytic models in support of the organization's clinical, operations and business initiatives and priorities.

May 2020–May 2021 **Graduate Research Assistant**, *The Ohio State University*, Columbus, OH, USA.
Department: College of Public Health; Supervisor: Elisabeth Root, Ayaz Hyder

- Led the development team (3 graduate students and 4 undergraduate students) to develop & deploy (including data architecture, software developments, translating requirements, infrastructure design, and implementation) the COVID-19 surveillance web tools for the Ohio Department of Health and over 300 schools in the central Ohio.

Aug 2016–May 2020 **Graduate Research Assistant**, *The Ohio State University*, Columbus, OH, USA.
Department: Integrated Systems Engineering; Lab: Security and Efficiency Through Analytics Laboratory; Supervisor: Theodore T. Allen

- Led the research team (4 graduate students) at OSU and collaborated with the tech team at ROSEN Group to explore new features in anomaly detection in the oil and gas pipelines with more interpretable and accurate AI modeling methods.
- Conducted the project "A Framework For Economical Cyber Security Inspection and Assurance" funded by National Science Foundation (NSF Grant #1912166) from May 2019 to Mar 2020.
- Conducted the project "Data-Driven Cyber Vulnerability Maintenance" funded by National Science Foundation (NSF Grant #1409214) from Aug 2016 to Nov 2018.
- Developed the cyber vulnerability risk prediction model with social media indications based on machine learning models. Worked with Pardee Bob and Geoffery Shoupp from Office of the Chief Information Officer (OCIO) Enterprise Security at the Ohio State University to deploy the proposed system as an API product.
- Co-authored with Dr. Theodore Allen, Dr. Sayak Roychowdhury, and Akshay K. Murali to publish two journal articles and two conference papers in the field of Reinforcement Learning & Discrete Events Simulation.
- Co-authored with Dr. Theodore Allen to write a book about machine learning applications in operations optimization and supply chain optimization.

Project Work

Sep 2020–May 2021 **Advanced AI in Pipeline Anomaly Classification**, *Funded by ROSEN Group*.
Team Leader: Enhao Liu; Team Members: 4 graduate students; Supervisor: Dr. Theodore Allen.

- Leveraged multiple image processing methods including wavelet transformation, saliency detection, contrast calculation, and blob detection methods to generate key features from millions of Magnetic Flux Leakage sensor data of the oil & gas pipelines.
- Achieved over 95% precision & recall for distinguishing pipeline defects in the proposed machine learning model with hyperparameters optimization.

Apr 2020–May 2021 **COVID-19 Comprehensive Monitoring Team at OSU**, *Funded by Ohio Department of Health*.
Lead developer: Enhao Liu; Team Members: 3 graduate students, 4 undergraduate students; Supervisor: Dr. Elisabeth Root, Dr. Ayaz Hyder.

- Led the full infrastructure of the R Shiny-based dashboard (over 50,000 code lines) of the COVID-19 surveillance for the Ohio Department of Health, including data architecture, software developments, translating requirements, infrastructure design, and implementation on statistical and epidemiological models (including EWMA, Relativ Risk, Space-Time Permutation Scan Statistics, etc.). Public-facing site is available: link.
- Developed a school-level COVID-19 surveillance tool called COVID-19 Analytics and Targeted Surveillance System (CATS) for over 20 school districts to facilitate superintendents to make learning mode schedules for over 300 schools, 180,000 students, and 20,000 staffs. One of the public-facing dashboards is available on the website of the Hilliard City School District: link

May 2019–Mar 2021	**Explainable AI Study: Optimal Classification Tree**, *NSF Grant #1912166*.

Graduate Research Assistant: Enhao Liu; Supervisor: Dr. Theodore Allen.

○ Developed the Python & R package for the optimal classification tree (OCT) algorithm based on mixed-integer linear programming, which supports Gurobi, CPLEX, and GLPK solvers.

○ Applied OCT to the defects classification in pipelines and showed that OCT is comparable to the ensemble learning tree models (such as XGBoost) in terms of model accuracy (96% precision & recall).

Jul 2019–Mar 2020	**Twitter-based Detection for Exploited Vulnerabilities**, *NSF Grant #1912166*.

Graduate Research Assistant: Enhao Liu; Supervisor: Dr. Theodore Allen.

○ Wrote scraping scripts using Python to collect recent five-year Twitter data (millions of tweets) and other online resources about 70,000 distinct software vulnerabilities spanned over 9,000 different vendors.

○ Developed feature generation and trained classification models to early detect exploited vulnerabilities based on XGBoost using scikit-learn framework.

Aug 2016–Nov 2018	**Cyber Vulnerabilities Maintenance Management**, *NSF Grant #1409214*.

Graduate Research Assistant: Enhao Liu; Supervisor: Dr. Theodore Allen.

○ Implemented data integration & analysis using R to process and statistically analyze 2.3 million scanned data about vulnerabilities, which provides the database for the development of cyber vulnerability maintenance policies.

○ Constructed a comprehensive framework based on POMDPs to derive cost-effective cyber vulnerability maintenance policies that not only significantly ($2x \sim 10x$) reduce expected maintenance expenditures per host but also address the incomplete inspection issue on muti-type hosts.

○ Worked with Dr. Theodore Allen and Dr. Sayak Roychowdhury to construct a new Monte Carlo-Bayesian Reinforcement Learning (MC-BRL) framework to address parameters uncertainty in estimating transition probabilities and expected costs of MDPs in which multiple identical-systems have been considered as well.

○ Collaborated with Dr. Theodore Allen and a graduate student Akshay K. Murali to build simulation models in ARENA to evaluate the cyber maintenance policies under different scenarios in terms of maintenance costs and security hardening.

Publications

Journal Articles

ASMBI 2019	**Liu, E.**, Allen, T. T., & Roychowdhury, S. (2019). Cyber vulnerability maintenance policies that address the incomplete nature of inspection. *Applied Stochastic Models in Business and Industry*, 35(6), 1390-1410.
CAIE 2018	Allen, T. T., Roychowdhury, S., & **Liu, E.** (2018). Reward-based Monte Carlo-Bayesian reinforcement learning for cyber preventive maintenance. *Computers & Industrial Engineering*, 126, 578-594.

Conference Papers

WSC 2019	Murali, A. K., **Liu, E.**, & Allen, T. T. (2019). Discrete event simulation of cyber maintenance policies according to nested birth and death processes. In *2019 Winter Simulation Conference (WSC)* (pp. 774-785). IEEE.
WSC 2018	Allen, T. T., & **Liu, E.** (2018). Forecasting cyber maintenance costs with improved scan analytics using simulation. In *2018 Winter Simulation Conference (WSC)* (pp. 1218-1225). IEEE.

Papers in Progress

AJE 2021	Kline, D., Hyder, A., **Liu, E.**, Rayo, M., Malloy, S., & Root, E. (2021). A Bayesian spatio-temporal nowcasting model for public health decision-making and surveillance. *American Journal of Epidemiology* (Accepted)
OR 2021	**Liu, E.,** & Allen, T. T. (2021). Optimal Classification Trees with Leaf-Branch-Interaction, Binary Split, and Node Complexity Constraints. *Operations Research*. (Submitted)

CAIE 2021	**Liu, E.,** & Allen, T. T. (2021). Twitter-based Cyber Security Warning and Prioritization System. *Computers & Industrial Engineering*. (Working)
Book 2021	Allen, T. T., & **Liu, E.** (2021). "Introduction to Machine Learning For Supply Chain Optimization - *Python, Statistics, and AI with Operations Analytics Industrial Datasets*." *Springer*. (Working)

Dissertations

Ph.D. Thesis 2021	**Liu, E.** (2021). "Innovative Simulation and Tree Models and Reinforcement Learning Methods with Applications in Cybersecurity". Ph.D. Thesis. The Ohio State University.
Master Thesis 2017	**Liu, E.** (2017). "Logistic Regression Model for Predicting Warning "Inciden" Rates and Implications for the Common Vulnerability Scoring System". Master Thesis. The Ohio State University.

Speaking

Invited Research Talks

2020	"Monitoring COVID-19 in Ohio & at Ohio State", Infectious Diseases Institute Virtual COVID-19 Symposium, Columbus, OH (Dec 03, 2020).
2020	"Cyber Security Threat Analysis Using Twitter Classification Models", INFORMS Annual Meeting – Session of Decisions About Information and Miss-Information in Social Media, Virtual (Nov 13, 2020).
2019	"Twitter-based Cyber Security Warning and Prioritization System", INFORMS Annual Meeting – Session of Information Operations in Social Networks, Seattle, WA (Oct 23, 2019).

Poster Sessions

2018	"Cyber Vulnerability Maintenance and Optimal Learning", Research Poster Session in the Department of Integrated Systems Engineering, OSU, Columbus, OH (Nov 2018).
2016	"Data-Driven Cyber Vulnerability Maintenance", Research Poster Session in the Department of Integrated Systems Engineering, OSU, Columbus, OH (Sep 2016).

Teaching

Spring 2020	**Teaching Assistant, ISE-5194 Group Studies in Integrated Systems Engineering (Introduction to Operations Analytics)**, *The Ohio State University*.
	○ Co-deigned the undergraduate & graduate course with Dr. Theodore Allen, with goal of teaching students to implement industrial applications with machine learning models in Python, with topics including supervised & unsupervised machine learning, cross-validation, sampling approach, regression, decision tree, ensemble learning trees, LSTM, hyperparameter optimization, Markov decision processes, and Q-learning.

Service

Journal Referee

Computers & Industrial Engineering

Academic Conference

2021	Session Chair, Artificial Intelligence for Social Media, 2021 INFORMS Annual Meeting

Service to University

2020	Member, COVIDCommons - a data platform to support COVID related research, The Ohio State University

External Service

2020	Member, InnovateOhio Platform Team
	○ Managed the data infrastructure for COVID-19 surveillance in the State of Ohio.