



YING WU COLLEGE
OF COMPUTING

RESEARCH REPORT 2023



Table of Contents

1	Dean's Message	30	Cybersecurity
2	Advanced Algorithms	34	Data Mining
6	Applied AI	38	Extended Reality
10	Big Data	43	Human Computer Interaction
15	Bioinformatics and Medical Informatics	46	Image and Video Processing
18	Blockchain Technologies	48	Machine Learning, Deep Learning and Reinforcement Learning
20	Cloud/Edge Computing and High-Performance Computing	53	Mobile Computing
24	Computing Education Research	55	Software Engineering and Programming Languages
26	Computer Graphics	58	Facts and Figures



Ying Wu College of Computing (YWCC) was established in 2001 within New Jersey Institute of Technology (NJIT) and consists of the computer science, data science and informatics departments. NJIT is a top 100 national university classified as a Carnegie R1 research university and YWCC has grown rapidly to become the second-largest college at NJIT - a trend we expect to continue in the near future. Enrolling more than 4,000 students at all levels and graduating more than 1,000 computing professionals every year, YWCC is the largest generator of computing tech talent in the greater New York metro area.

The tenure-track faculty and Ph.D. student population of YWCC has also grown significantly over the past five years, and these talented individuals are responsible for the high-quality academic research described in this report. Their research is funded by a variety of government agencies and corporate entities, published in top venues, and in some cases patented and commercialized.

Spanning a wide spectrum of topics, from human-computer interaction to cybersecurity to sophisticated data science algorithms, the research conducted by YWCC ranges from deep mathematical theory to practical applications. Committed to sharing beyond publication, much of the software developed in these projects is made available to the general scientific community through open-source repositories.

I invite you to read through the abstracts of the research projects described in this report and encourage you to reach out to the individual researchers for more details, if needed. We welcome new ideas, collaborations and any form of research partnership imaginable.

Sincerely,

A handwritten signature in black ink that reads "C. Gotsman". The signature is fluid and cursive, with a long horizontal line extending from the end of the name.

Craig Gotsman
Distinguished Professor
Former Dean, Ying Wu College of Computing
New Jersey Institute of Technology

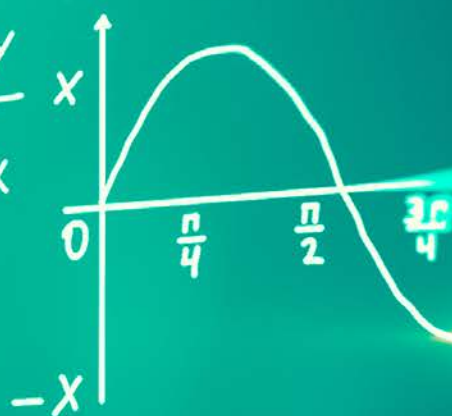
Advanced Algorithms

Our research is focused on algorithms with formal guarantees of performance. These include continuous optimization, statistically robust risk analysis, algorithms for dynamically evolving graphs, and approximation algorithms for various graph problems with concrete industrial applications.

$$\sqrt{\sum_{t=2}^n (y_t - \bar{y}_1)^2 \cdot \sum_{t=2}^n (y_{t-1} - \bar{y}_2)^2}$$

$$\varepsilon) = \tilde{S}^2(\varepsilon) = \frac{\sum_{i=1}^n e_i^2}{n-2} \cdot (1) y_x * \frac{S_y}{S_x} x$$

$$= \frac{\sum_{t=2}^n y_t}{n-1}; \quad \bar{y}_2 = \frac{\sum_{t=2}^n y_{t-1}}{n-1};$$



$$= \frac{dQ_{ex}}{de} \cdot \frac{e}{Q_{ex}}; \quad \varepsilon_{im} = \frac{dQ_{im}}{de} \cdot \frac{e}{Q_{im}} \cdot \sqrt{\frac{q-3}{8/5}} \int \int \sqrt{x+\sqrt{y}} \, dx \, dy$$

$$e) = Q_{ex}(e) - e Q_{im}(e), \quad \text{Integrate}[1/(x^4 6 + x^2 2 x + 1), x]$$

$$= \frac{dQ_{ex}}{de} \Delta e - e \frac{dQ_{im}}{de} \Delta e - e Q_{im}, \quad (4) \quad \frac{8}{105} (x + \sqrt{y})^{5/2} (-2x + \sqrt{y})$$

$$b) = \int_0^1 (1-x)^{b-1} d\frac{x^a}{a} = \beta_{yx} = r \frac{1}{56} \left(7 + \sqrt{7(-5 + 4\sqrt{2})} \right)$$

$$\frac{(1-x)^{b-1}}{a} \Big|_0^1 + \frac{b-1}{a} \int_0^1 x^a (1-x)^{b-2} dx = f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

$$= \frac{1}{a} \int_0^1 x^{a-1} (1-x)^{b-2} dx - \frac{b-1}{a} \int_0^1 x^{a-1} (1-x)^{b-1} dx =$$

$$= \frac{1}{a} B(a, b-1) - \frac{b-1}{a} B(a, b), \quad r(\nabla_{X_f}, \nabla_{Y_f}) =$$

$$b) = \frac{b-1}{a+b-1} B(a, b-1) = r_{yx} * \frac{S_y}{S_x}$$



James Calvin

Professor
james.m.calvin@njit.edu

Research Areas: Global optimization, probabilistic analysis of algorithms

Global Optimization

To solve the pervasive optimization problems in engineering, science and commerce, we are developing “global optimization” algorithms where the objective is to solve optimization problems without getting stuck in local minima. This has applications in the design of fuel-efficient aircraft, the error rate of classification algorithms and financial investing.



Ioannis Koutis

Associate Professor
ioannis.koutis@njit.edu

Research Areas: Fast Linear system solvers, spectral graph algorithms, machine learning, data mining

Spectral Graph Algorithms

Graphs are objects of central interest in data science. A graph can be mapped to linear operators whose spectral properties encode connectivity information, enabling the design of numerical algorithms for various problems on graphs. The practical applicability of such algorithms hinges on the existence of fast solvers for fundamental computational problems, such as systems of linear equations and other generalized regression problems. We have designed such solvers and leveraged them in the design of new graph algorithms based on spectral graph theory for graphs with a prescribed cut structure. As a concrete application, we developed new algorithms for hypergraph partitioning, a key problem in Electronic Design Automation. Our algorithms broke previous records on multiple benchmarks and received the best paper award at the 41st ACM/IEEE ICCAD, in 2022.

Exact Algorithms

Multiple well-studied computational problems are conjectured to require exponential time for their solution. Current research aims to develop a detailed understanding of the computational complexity of these problems beyond the classical NP-completeness theory. The design of faster exact algorithms for such problems and their parameterized versions is of key importance in the area. In this context, we pioneered the general method of algebraic fingerprints that reduces various combinatorial problems to monomial detection problems that are, in turn, solved via algebraic algorithms. This research led to breakthrough results for classical algorithmic problems, such as the Hamiltonian cycle problem and single exponential time algorithms for problems parameterized by treewidth. Among these breakthrough algorithms, our work on Directed Hamiltonicity and Outbranching problems received the best paper award at the 44th ICALP in 2017.



Marvin Nakayama

Professor
marvin@njit.edu

Research Areas: Monte Carlo simulation, risk analysis, applied probability, statistics

Efficient Computational Methods for Risk Assessment

Many disparate fields suffer from uncertainties with detrimental consequences, such as large losses in financial portfolios or failures of critical infrastructure due to natural disasters. Modern society crucially depends on gaining a better understanding of the likelihood and impacts of such calamitous rare events. Our work devises novel approaches for substantially reducing statistical errors in Monte Carlo simulation, a computational technique that can be employed to study risks in decision-making and analytics. One project focuses on designing improved methods for probabilistic safety assessments of nuclear power plants. The work also applies to evaluating the reliability of complex systems such as aircraft navigation computers, package-tracking systems for overnight delivery companies and the dependability of supply chains.



Baruch Schieber

Professor
baruch.m.schieber@njit.edu

Research Areas: Algorithms, mathematics of artificial intelligence (AI), optimization

Dynamic Graph Algorithms

Graphs are used extensively to model various kinds of networks, such as transportation or social networks. In most real-life applications these networks change over time so their characteristics are changing as well. The goal of dynamic graph algorithms is to compute these characteristics over time as efficiently as possible. The required output can be recomputed from scratch at each time point, however, in many cases, the slow pace of change relative to the size of the network enables much faster computation. One such characteristic is the maximal independent set (MIS) of the graph, which has extensive connections to many fundamental combinatorial optimization problems, such as maximum matching, minimum vertex cover and graph coloring. We developed several dynamic algorithms for computing MIS including the first sublinear amortized update time algorithm for maintaining an MIS in dynamic graphs.

Location Problems on Euclidean Metrics

Location problems are an important class of combinatorial optimization problems that arise in applications such as choosing facility sites in a supply chain, placing servers in a telecommunication network and clustering data. The underlying distance function in many cases is Euclidean. It is natural to ask whether the Euclidean metrics can be leveraged to obtain more efficient algorithms than the ones known on a general metric space. We considered one such location problem, the classical k-supplier problem and showed that there exists an algorithm for this problem on Euclidean metrics that beats the lower bound on the time required for any such algorithm on a general metric space.



Pan Xu

Assistant Professor
pan.xu@njit.edu

Research Areas: Artificial intelligence, approximation and randomized algorithm, ridershare and crowdsourcing markets

Algorithm Design to Combat Discriminations in Rideshare Services

Rideshare platforms, such as Uber and Lyft, have gained increasing popularity in recent years. One of the central tasks facing Uber and Lyft is the matching policy pairing drivers and riders. Recently, it has been reported that drivers cancel riders based on their demographic attributes such as gender, race and disability, either intentionally or unintentionally. In our project, we try to leverage the power of algorithm design to curb discriminative cancellations from drivers to riders and improve the social welfare overall.

Applied AI

Artificial Intelligence (AI) has been applied in a dizzying array of applications transforming our society with significant impacts across domains, including marketing, healthcare, social good applications, entertainment, etc. Advancing the core technologies with cutting-edge solutions is our strength, demonstrated through consistent and crucial contributions published at leading venues in the field, such as the AAAI Conference, and via partnerships with industry for technology transfer, such as NEC, Accenture, Adobe and Qualcomm.





Cristian Borcea

Professor
borcea@njit.edu

Research Areas: Mobile computing and sensing, machine learning for mobility, predictive computational advertising

Predictive Online Advertisement

Online advertising is a multi-billion-dollar industry. Publishers typically sell ad impressions that appear in web pages through auctions held in ad exchanges in real-time. A publisher accepts the winning bid if it is higher than a given reserve price for an ad impression. We created a machine learning model that determines the optimal reserve prices for individual ad impressions in real-time, using a multi-task learning framework to predict the lower bounds of the highest bids. The revenue from the ads can decrease drastically if users employ ad-blockers in the browsers. To mitigate this problem, we created a personalized counter ad-blocking system that dynamically chooses a counter ad-blocking strategy, based on the predicted individual user's probability to whitelist the page or site they intend to view in the ad-blocker. This strategy increases the revenue, while minimizing the user annoyance associated with showing too many ads.



Mark Cartwright

Assistant Professor
mark.cartwright@njit.edu

Research Areas: Machine listening, interactive machine learning, human-computer interaction, audio processing, music information retrieval

Machine Listening with Limited Labeled Data

Progress in machine listening has historically been hindered by limited amounts of labeled data. This is due to a variety of reasons: we don't share and tag audio as we do images; licensing terms prevent the distribution of annotated musical audio; and, unlike with images, the average person cannot effectively label audio with just a glance and a click — it takes time to listen. To address this, we have developed and studied methods for self-supervised audio representation learning, few-shot sound event detection, crowdsourced audio annotation and active learning.

Interactive Open-World Machine Listening

Smart acoustic sensing powered by machine listening has the potential to provide valuable information for numerous tasks, including understanding and improving the health of our cities (e.g., monitoring and mitigating noise pollution) and natural environments (e.g., monitoring and conserving biodiversity). However, contemporary machine listening models, such as sound event detection models, can only detect a small fraction of the sounds we care about in the world. To the "ears" of these models, unknown sound classes do not exist or are confused with known classes. This results in a limited view of the acoustic world by sound event detection models that may not align to the goals of end users, hindering the machine listening's transformative potential. We aim to develop methods and tools to detect, discover, describe and define "unknown" sound events in support of large-scale longitudinal audio analysis and model development.



Hai Phan

Assistant Professor
hai.n.phan@njit.edu

Research Areas: Trustworthy machine learning, data privacy and security, social good applications

Preserving Sensitive Health Care Data

Networks that share electronic health records containing a patient's personal and clinical data promise improved continuity of care and better health outcomes. However, these networks put highly sensitive patient information at risk and expose healthcare providers to legal jeopardy. We created DeepPrivate, a system that uses machine learning techniques to protect personal health information against cyber attacks.

Ecosystem for Federated Learning Mobile Applications

The challenges for federated learning (FL) systems on phones are concurrent management of multiple FL activities under resource constraints and frequent disconnections due to networking and battery issues. In addition, the system needs to provide rigorous privacy protection while retaining high model utility in real-world deployments. To address these challenges, we develop a novel FL system architecture integrated with scalable privacy-preserving solutions. This project is an impactful step to unlocking the power of millions of mobile devices through federated machine-learning applications while preserving user privacy.



Guiling "Grace" Wang

Distinguished Professor
gwang@njit.edu

Research Areas: Deep learning, FinTech, blockchain technologies, intelligent transportation systems

Financial Market Prediction by Leveraging Deep Learning

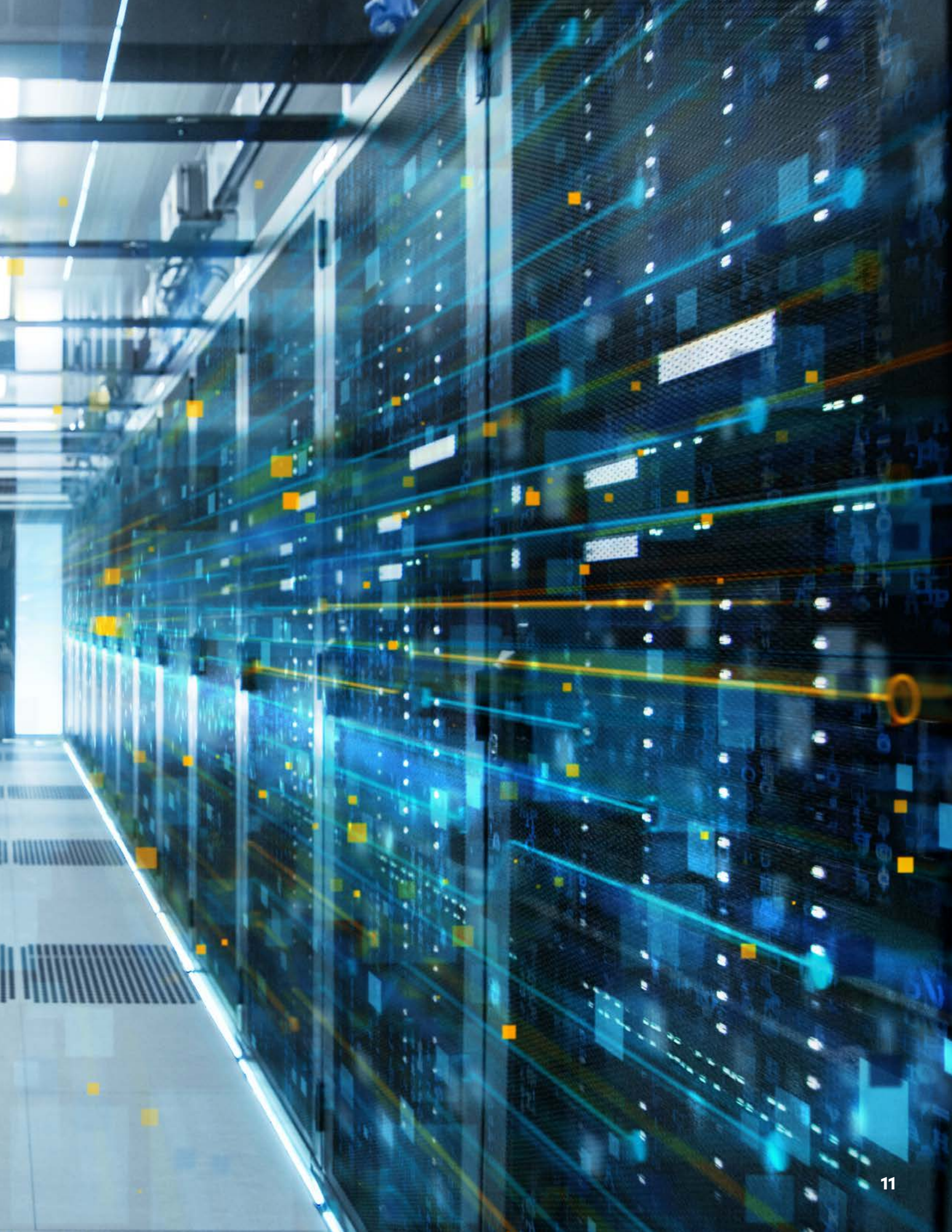
Finance market prediction attracts much attention from both industry and academia. Despite significant efforts, the results are unsatisfactory due to the inherent stochasticity nature of the finance market driven by various unpredictable factors, such as the emotion of market participants and macroeconomic news. To address the problem, we employ sentiment analysis on the news to learn public emotion and develop a deep learning model to predict the stock price. We also incorporate the volatility index to investigate market sentiment further and utilize Wasserstein Generative Adversarial Network (WGAN) for multi-step movement prediction. The generated sequence from GAN with news context learning and prior noise can simulate the future price pattern. Extensive experiments on multiple stock indices worldwide illustrate that our model achieves consistent and significant improvements over multiple popular benchmarks on both interval and point prediction.





Big Data

We work on a variety of big data topics, including high performance data analytics, massive-scale graphs, dimension reduction, complex workflows, big data management and computing, and green computing and networking, as well as data mining research with emphasis on keyword search within graphs and trees. The research is a combination of algorithms and techniques to improve performance of large-scale data systems and extract relevant information. Our research has been presented at top tier conferences, including VLDB, SIGMOD, WWW, KDD, IPDPS, SC, INFOCOM, and eScience, and received various awards including the SIGMOD Test of Time Award in 2015.





David Bader

Distinguished Professor
bader@njit.edu

Research Areas: Data science, high-performance computing

High Performance Algorithms for Interactive Data Science at Scale

A real-world challenge in data science is to develop interactive methods for quickly analyzing new and novel data sets that are potentially of massive scale. This project will design and implement fundamental algorithms for high performance computing solutions that enable the interactive large-scale data analysis of massive data sets. Based on the widely-used data types and structures of strings, sets, matrices and graphs, this methodology will produce efficient and scalable software for three classes of fundamental algorithms that will drastically improve the performance on a wide range of real-world queries or directly realize frequent queries. These innovations will allow the broad community to move massive-scale data exploration from time-consuming batch processing to interactive analyses that give a data analyst the ability to comprehensively, deeply and efficiently explore the insights and science in real world data sets. By enabling the increasing number of developers to easily manipulate large data sets, this will greatly enlarge the data science community and find much broader use in new communities.

Building Faster, Energy-Efficient Analytics Pipelines for Decision-Making

Big data analysis is used for problems related to massive data sets. Today, these data sets are loaded from storage into memory, manipulated and analyzed using high performance computing (HPC) algorithms and then returned in a useful format. This end-to-end workflow provides an excellent platform for forensic analysis; however, there is a critical need for systems that support decision-making with a continuous workflow. HPC systems must focus on ingesting data streams; incorporating new microprocessors and custom data science accelerators that assist with loading and transforming data; and accelerating performance by moving key data science tasks and solutions from software to hardware. These workflows must be energy-efficient and easy to program, while reducing transaction times by orders of magnitude. Analysts and data scientists must be able to ask queries in their subject domain and receive rapid solutions that execute efficiently, rather than requiring sophisticated programming expertise.

Scalable Graph Algorithms

Our research is supported in part by an NVIDIA AI Lab (NVAL) award. NVIDIA makes graphics processing unit (GPU) accelerators such as the DGX Deep Learning server. We contribute to RAPIDS.ai, an open GPU data science framework for accelerating end-to-end data science and analytics pipelines entirely on GPUs. These new analytics pipelines are more energy-efficient and run significantly faster, which is critical for making swift, data driven decisions.



Vincent Oria

Professor
vincent.oria@njit.edu

Research Areas: Multimedia databases, spatio-temporal databases, recommender systems

Dimensionality and Scalability Issues in High-Dimensional Spaces

When researching fundamental operations in areas such as search and retrieval, data mining, machine learning, multimedia, recommendation systems and bioinformatics, the efficiency and effectiveness of implementations depends crucially on the interplay between measures of data similarity and the features by which data objects are represented. When the number of features known as data dimensionality is high, the discriminative ability of similarity measures diminishes to the point where methods that depend on them lose their effectiveness. Our research looks at the interplay between local features, the intrinsic dimensionality and their application to search, indexing and machine learning.

Vincent Oria

Multi-Instrument Database of Solar Flares

Solar flares are the most prominent manifestation of the sun's magnetic activity. They emit radiation that could potentially damage power systems, interfere with civilian and military radio frequencies and disrupt spacecraft operations. To improve analysis, in collaboration with the department of physics, we aim to integrate, clean, and enrich solar data captured by various solar flare observing instruments around the world and are using them for some predictive analysis tasks.



Senjuti Basu Roy

Associate Professor
senjuti.basuroy@njit.edu

Research Areas: Human-in-the-loop large-scale data analytics, optimization algorithms

Big Data Analytics Laboratory - Data Analytics with Humans-in-the-loop

The Big Data Analytics Lab (BDaL) is an interdisciplinary research laboratory that focuses on large-scale data analytics problems rising in different application domains and disciplines. One focus of our lab is to investigate an alternative computational paradigm that involves humans-in-the-loop for big data. These problems arise at different stages in a traditional data science pipeline, such as data cleaning, query answering, ad-hoc data exploration or predictive modeling, as well as from emerging applications. We study optimization opportunities that arise because of this unique man-machine collaboration and address data management and computational challenges. Our focus application domains are crowdsourcing, social networks, health care, climate science, retail and business, naval applications and spatial data.



Chase Wu

Professor
chase.wu@njit.edu

Research Areas: Big data, machine learning, green computing and networking, parallel and distributed computing

Revolutionizing Big Data Scientific Computations

Next-generation scientific applications are experiencing a rapid transition from traditional experiment-based methodologies to large-scale simulations featuring complex numerical modeling with a large number of tunable parameters. Such model-based simulations generate colossal amounts of data, which are then processed and analyzed against experimental or observation data for parameter calibration and model validation. The sheer volume and complexity of such data, the large model-parameter space and the intensive computation make it practically infeasible for domain experts to manually configure and tune hyperparameters for accurate modeling in complex and distributed computing environments. We develop visualization algorithms for 3D volume data generated by scientific computations on supercomputers and apply machine learning techniques to automate, expedite and optimize the parameter tuning process in model development.

Modeling and Optimizing Big Data Ecosystems

The execution of big data workflows is now commonly supported on reliable and scalable data storage and computing platforms such as Hadoop. There are a variety of factors affecting workflow performance across multiple layers in the technology stack of big data ecosystems. Modeling and optimizing the performance of big data workflows is challenging because the compound effects of such technology layers are complex and opaque to end users. We develop a cross-layer coupled design framework, which integrates information theory-based feature selection and stochastic approximation-based profiling to automate and optimize the configuration of big data ecosystems.

Chase Wu

Optimizing Distributed Training and Inference of Deep Neural Networks (DNNs)

Deep Neural Networks (DNNs) have grown rapidly in size and complexity, requiring various data/model/tensor parallelization techniques to make training/inference practically feasible. For example, BLOOM 176B and Megatron-Turing 530B require terabytes of memory and zettaflops of compute. We represent parallelized DNNs as workflows and develop new approaches to workflow partitioning, mapping, and scheduling alongside memory saving techniques such as activation recomputation to optimize the training and inference processes of DNNs in heterogeneous multi-node, multi-GPU/CPU systems.

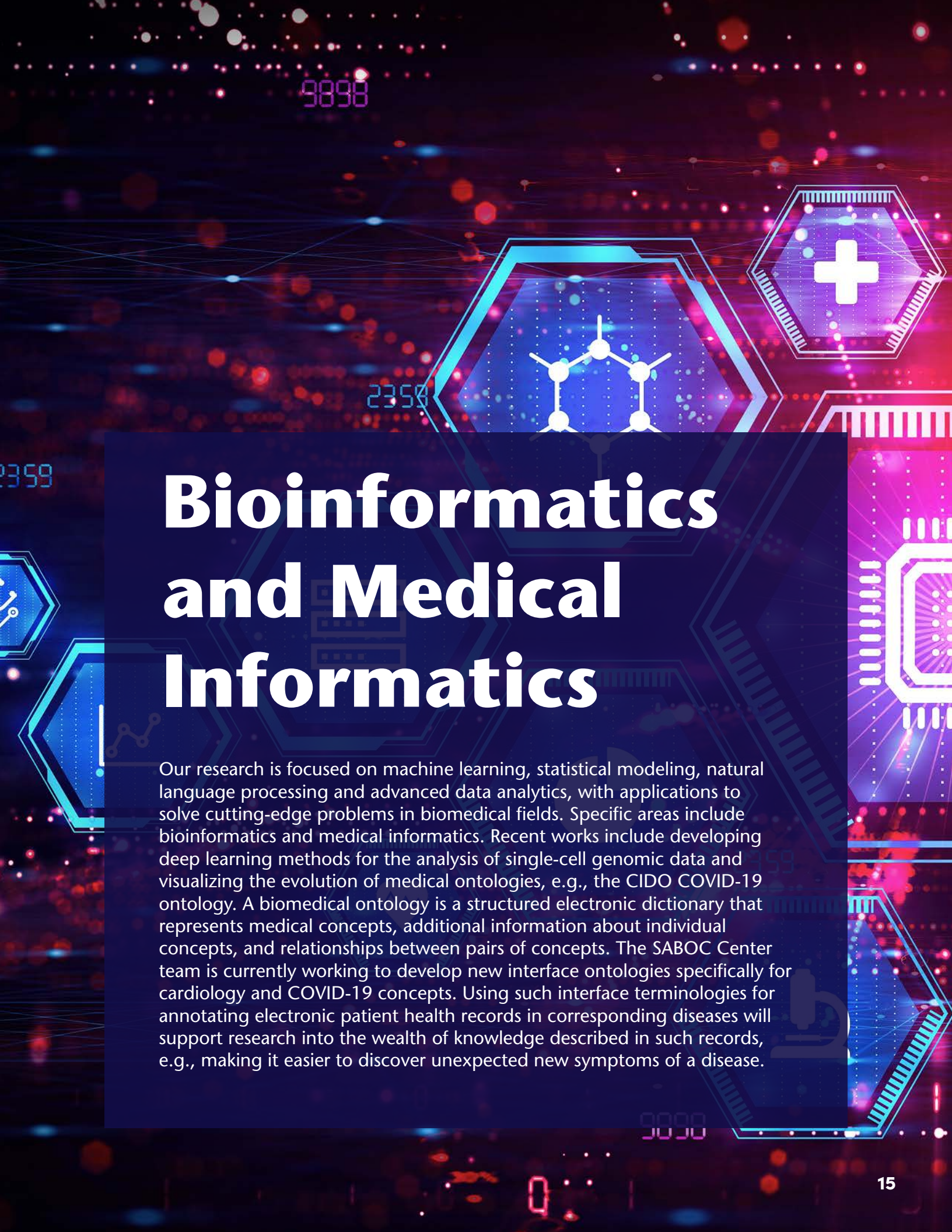
Reducing Energy Consumption in Big Data Computation

The transfer of big data across high-performance networks consumes a significant amount of energy. Employing two widely adopted power models — power-down and speed scaling — we have made inroads into green computing and networking in big data environments. Our approach allows network providers to reduce operational costs and reduce carbon dioxide emissions.

Uncovering Low-Level, Hazardous Radiation

Radioactive substances and biological agents present a serious threat to public health and safety, particularly in densely populated areas. Through the collection and analysis of large amounts of sensor measurements, we develop reliable tools to detect and contain radioactive materials to protect the populace and reduce the risk of radiological dispersal devices, such as so-called dirty bombs.





Bioinformatics and Medical Informatics

Our research is focused on machine learning, statistical modeling, natural language processing and advanced data analytics, with applications to solve cutting-edge problems in biomedical fields. Specific areas include bioinformatics and medical informatics. Recent works include developing deep learning methods for the analysis of single-cell genomic data and visualizing the evolution of medical ontologies, e.g., the CIDO COVID-19 ontology. A biomedical ontology is a structured electronic dictionary that represents medical concepts, additional information about individual concepts, and relationships between pairs of concepts. The SABOC Center team is currently working to develop new interface ontologies specifically for cardiology and COVID-19 concepts. Using such interface terminologies for annotating electronic patient health records in corresponding diseases will support research into the wealth of knowledge described in such records, e.g., making it easier to discover unexpected new symptoms of a disease.



Usman Roshan

Associate Professor
usman.w.roshan@njit.edu

Research Areas: Machine learning, medical informatics

Deep Learning for Medical AI Problems

We work on several medical AI problems involving images and DNA. In collaboration with neuroscientists, we are creating new models for tumor identification in brain MRI images and are studying them across different data sources. With vascular surgeons we are proposing novel models for identifying vessel and plaque in vascular ultrasound images from real patients. We are studying simple random networks for the classification of histopathology slide images and find them to be highly accurate there.



Zhi Wei

Professor
zhiwei@njit.edu

Research Areas: Machine learning, statistical modeling, bioinformatics

Deep Learning Methods to Integrate Biological Information for Analysis of Single-cell RNAseq Data

The broad long-term objective of the project concerns the development of novel machine learning methods and computational tools for modeling genomic data motivated by important biological questions and experiments. The analysis of single-cell RNAseq (scRNAseq) data presents substantial computational and bioinformatics challenges. The specific aim of the project is to develop novel model-based deep learning methods with prior biological information considered for modeling scRNAseq data. These problems are all motivated by the collaborations with biomedical investigators. The proposed approaches are designed to integrate biological information for improving both analytical performance and biological interpretability. The methods hinge on novel integration of biological insights and deep learning methods for analysis of the noisy, sparse, and over-dispersed scRNAseq data, including zero-inflated negative binomial model, autoencoder, deep embedding, hyperbolic embedding and reversed graph embedding.



James Geller

Professor
james.geller@njit.edu



Michael Halper

Professor
michael.halper@njit.edu



Yehoshua Perl

Professor
yehoshua.perl@njit.edu

James' Research Areas: Medical informatics, medical ontologies and terminologies, applied social network analysis, DEI (diversity, equity and inclusion)

Michael's Research Areas: Controlled terminologies, ontologies, object-oriented databases, conceptual modeling, open data models

Yehoshua's Research Areas: Medical informatics, semantic web, ontologies

Summarizing, Visualizing and Correcting Large Collections of Medical Terms

Biomedical professions have collected large ontologies of medical terms covering diseases, diagnosis, drugs, anatomy, bacteria, genes, chemicals, medical procedures and more. We developed a theoretical framework to create concise summaries of large ontologies and software tools to visualize these summaries. Using this software, medical experts can easily browse large medical ontologies, making it easier to recognize errors in structure and content. Recently, we developed machine learning methods for updating and curating biomedical ontologies. We are also developing methods to interpret large collections of formal medical knowledge in ontologies, which will aid in the prediction of dangerous drug interactions.



The background is a dark, abstract composition featuring several glowing, translucent spheres in shades of purple and blue. These spheres are interconnected by a network of lines and are surrounded by a dense field of binary code (0s and 1s) in a light blue or white color. The overall effect is a sense of digital connectivity and data flow.

Blockchain Technologies

Blockchain technology has attracted increasing attention from both academia and industry. Its popular applications include cryptocurrency, smart contracts, non-fungible tokens (NFT), Web3 and many others. It is an interdisciplinary field encompassing distributed systems, security, privacy, cryptography, database and compiler. We conduct fundamental and applied research in blockchain consensus and decentralized applications. Our projects are funded by federal agencies, such as FHWA, and the results of our research are published in top venues, such as IEEE Transactions on Parallel and Distributed Systems and IEEE ICDCS.



Guiling "Grace" Wang

Distinguished Professor
gwang@njit.edu

Research Areas: Deep learning, FinTech, blockchain technologies, intelligent transportation systems

Decentralized Vehicle Credential Management System Based on Consortium Blockchain

To realize the full potential of connected and automated vehicles to transform the US transportation system, it is important to ensure the vehicle communication is secure and privacy-preserving. In this situation, a digital certificate is critical to provide communication integrity, authenticity and privacy. Existing solutions based on centralized security credential management systems are inherently vulnerable to emerging cyber-security risk and single-point failure. To address the weakness of centralized systems being adopted, this project seeks to utilize the recent breakthrough of blockchain technology, in particular the consortium blockchain, to design and implement an innovative decentralized vehicle credential management system. The proposed decentralized credential management system based on consortium blockchain can not only greatly improve the robustness and security of the credential management system for vehicles and thus realizes high-security assurance for authentic V2V and V2I communication, but also incorporate state-of-the-art achievements in cryptography to ensure strong anonymity to prevent hackers from tracing the locations of vehicles and ensuring vehicle privacy. This project was one of the four awardees among 122 proposals submitted to FHWA.

Blockchain-Based P2P Content Delivery

Peer-to-peer (P2P) content delivery provides many benefits compared with centralized content delivery networks (CDNs) and is complementary to popular decentralized storage networks such as Filecoin. However, reliable P2P delivery demands proper enforcement of delivery fairness. Still, most existing studies on delivery fairness are based on non-cooperative game-theoretic assumptions that are arguably unrealistic in the ad-hoc P2P setting. We propose an expressive yet minimalist security requirement for desired fair P2P content delivery and give two efficient blockchain-enabled and monetary-incentivized solutions FairDownload and FairStream, for P2P downloading and P2P streaming scenarios, respectively. Our designs not only ensure delivery fairness where deliverers are paid (nearly) proportional to their in-time delivery but also guarantee exchange fairness where content consumers and content providers are fairly treated.

(This is a collaborative work of Dr. Wang and Dr. Wu.)



Chase Wu

Professor
chase.wu@njit.edu

Research Areas: Big data, machine learning, green computing and networking, parallel and distributed computing

A Blockchain-Based Decentralized IIoT Management System

Blockchain has been increasingly used to secure data collection, storage and transfer in various Industrial Internet of Things (IIoT) environments. We design a blockchain-based decentralized IIoT management system for censorship resistance, which employs a diffusion mechanism to deliver messages from sensors to full nodes, an augmented consensus protocol to check data losses and facilitate opportunistic outcome delivery, and a public key-based aggregation method to reduce communication complexity and signature verification.

Cloud/Edge Computing and High-Performance

Clouds and edges form the backbone of modern computing and are fast evolving. The research explores multiple aspects of cloud/edge computing, including architecture, system software, algorithm and software system designs. One key strength of the research is enabling applications to fully leverage the computing capability of hardware and devices to achieve optimal data processing speeds. The research results will benefit many application domains and scenarios (e.g., AI and big data analytics, cyber-physical systems, scientific computing and interactive services) and have attracted interest from both industry and open-source communities.





Xiaoning Ding

Associate Professor
xiaoning.ding@njit.edu

Research Areas: Cloud/edge computing infrastructures, system designs for AI/DL, system software designs, database and data storage systems

Management of Disaggregated and Dynamic Resources in Clouds and Edges

In the post Moore's law era, computing resources are undergoing fundamental changes in many aspects (e.g., types, architectures, and features). In clouds and edges, computing resources are increasingly heterogeneous (e.g., many varieties of processors and accelerators), disaggregated (e.g., local and remote memory pooled together and made available through fast network) and dynamic (e.g., resource availability changing over time). These changes enable new computing paradigms and optimization opportunities, yet raise new challenges in resource management.

System Software for Scalable Computation in the Cloud

As computational resources continue to increase, we need ways to scale the performance of these computers by taking advantage of the extra resources. The objective is to guarantee that applications in the cloud can achieve higher performance when presented with more resources.



Alex Gerbessiotis

Associate Professor
alexg@njit.edu

Research Areas: Architecture-independent parallel algorithm design and implementation

Multi-Core and Many-Core Algorithm Design, Analysis and Implementation

We study models of computation that abstract and capture parallelism in the presence of multiple memory hierarchies and cores. New approaches are needed to make multi-core architectures accessible to software designers in domains such as machine learning and big data. Abstracting the programming requirements of such architectures in a useful and usable manner is necessary to increase processing speed and improve memory performance.

Parallel Computing Techniques in Sequential Serial Computing

The norm in computing is to port sequential algorithms that work on one processor into multi-core or parallel algorithms intended for multiple cores and processors. Amdahl's Law highlights the limitations of using multiple cores in programs with an inherently sequential component that is not amenable to parallelization. We address this by exploring the utilization of parallel computing techniques to speed up a sequential program by exploiting the multiple memory hierarchies present in contemporary microprocessors, even if its multi-core capabilities are left unexploited.



Jing Li

Assistant Professor
jingli@njit.edu

Research Areas: Real-time systems, parallel computing, cyber-physical systems and reinforcement learning for system design and optimization.

Parallel Real-Time Systems

Real-time systems need to provide timing guarantees for latency-critical applications in cyber-physical systems that interact with humans or the physical environment. Examples span autonomous vehicles, drones, avionic systems and robotics to structural health monitoring systems and hybrid simulation systems in earthquake engineering. However, as parallel machines become ubiquitous, we face challenges in designing real-time systems that can fully utilize the efficiencies of parallel and heterogeneous computing platforms. We are developing parallel real-time systems by exploiting the untapped efficiencies in the parallel platforms, drastically improving the system performance of a cyber-physical system.

Scheduling for Interactive Cloud Services

Delivering consistent interactive latencies, such as response delays, is the key performance metric of interactive cloud services that significantly impacts user experience. The need to guarantee low-service latency, while supporting increasing computational demands due to complex functions of the services, requires parallel scheduling infrastructure to effectively harness parallelism in the computation and efficiently utilize system resources. Our research designs, analyzes and implements scheduling strategies that are measurably good and practically efficient to provide various quality-of-service guarantees on cloud service latency.



Andrew Sohn

Associate Professor
andrew.sohn@njit.edu

Research Areas: Adaptive computing infrastructure, autonomous migration of Linux virtual machines, maximizing cluster utilization

Enabling High-Performance Cloud Computing

The persistent uploading, downloading and processing of images, videos and files to and from the cloud can lead to inefficiency and delayed response times due to irregular computing demands. Our research focuses on live migration of virtual machines, as well as containers that will help alleviate the problem and improve cloud servers such as Amazon's Elastic Compute Cloud and Microsoft Azure. It will also help meet the power and computing requirements of mobile and enterprise cloud applications.

Scalable Parallel Graph Partitioning for Enabling Real-Time Analytics

We are working on high performance computing for large-scale data, in particular largescale graph partitioning projects called HARP and S-HARP (scalable HARP) designed and implemented with collaborators at the NASA Ames Research Center and the Lawrence Berkeley National Laboratory. Large-scale graph partitioning is critical in real-time social network analytics and is particularly challenging when dealing with dynamic graphs that change over time, as there needs to be balance of partition quality and execution time. We established a framework for partitioning dynamic graphs for NASA applications and continue to improve the technology for real-time social network analytics on a cluster of personal computers.



Computing Education Research

Computing education research (CER) makes discoveries about teaching and learning about computing (broadly defined), and invents new ways for such teaching and learning to occur. To do so, computing education researchers often create educational technologies to support the learning and teaching of computing, focusing on how people can benefit pedagogically from using these interventions. CER is an example of discipline-based education research, which is part of the broader field of education and learning science research.



Michael Lee

Associate Professor
michael.j.lee@njit.edu

Research Areas: Human-computer interaction (HCI), computing education research (CER)

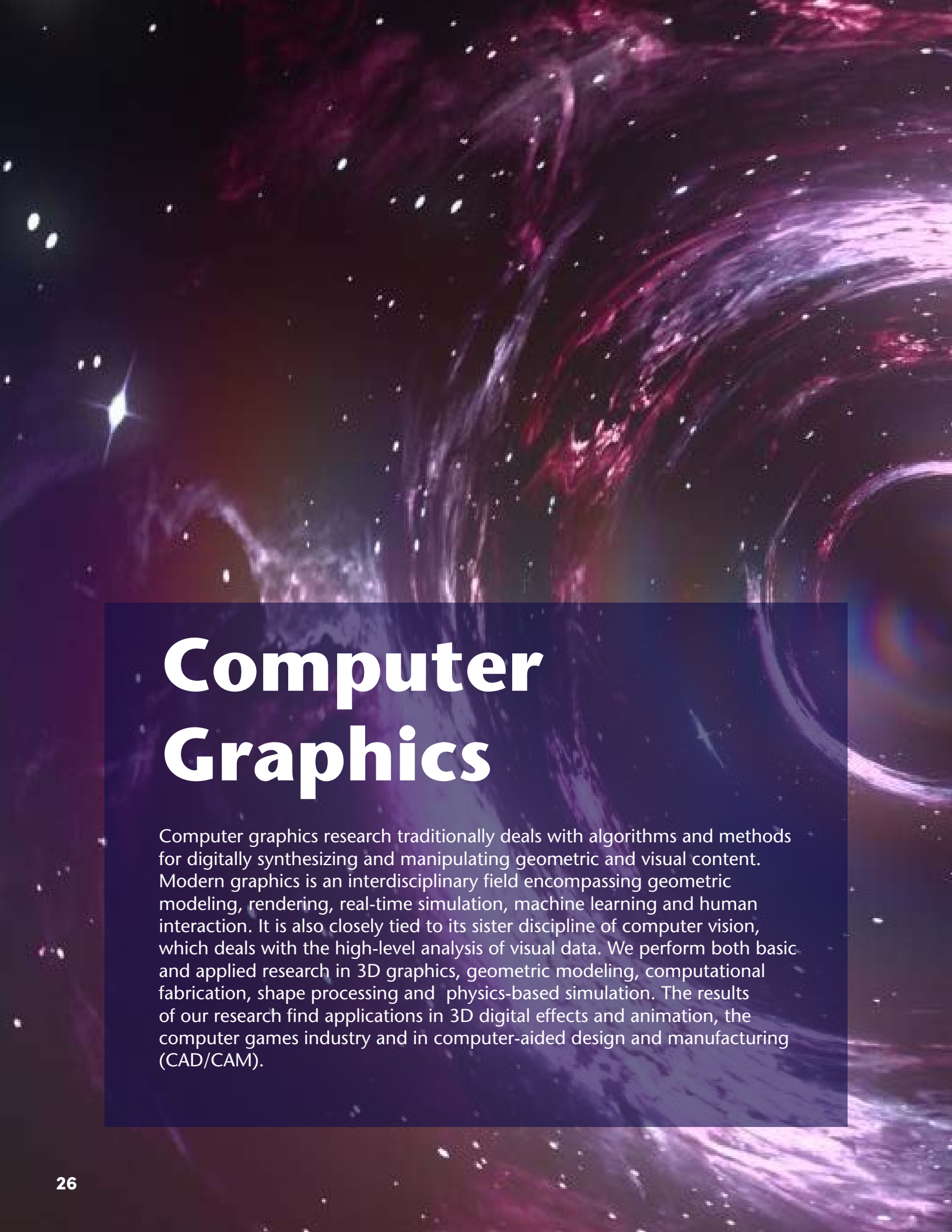
Advancing STEM Education with Gaming and Virtual Reality

We explore unique approaches to teaching STEM topics in formal and informal learning environments. Our work on Gidget—an online game (helpgidget.org) to teach students introductory programming concepts—is effective in attracting and engaging a broad audience, including women and underrepresented minority groups in computing. Our work with CSpresso uses an interactive VR environment to teach middle school and high school students introductory computer science concepts such as binary counting and sorting algorithms.

Increasing Diversity in STEM Through Mentorship

This project explores how to increase participation and diversity in computing, especially for underserved and underrepresented minorities in STEM. We partner with local nonprofits and schools to provide programming experience to K-12 students, specifically using near-peer mentors to teach and engage middle school and high school students in a programming camp called Newark Kids Code. We also provide an introductory computing course for in-service high school teachers during the summer.





Computer Graphics

Computer graphics research traditionally deals with algorithms and methods for digitally synthesizing and manipulating geometric and visual content. Modern graphics is an interdisciplinary field encompassing geometric modeling, rendering, real-time simulation, machine learning and human interaction. It is also closely tied to its sister discipline of computer vision, which deals with the high-level analysis of visual data. We perform both basic and applied research in 3D graphics, geometric modeling, computational fabrication, shape processing and physics-based simulation. The results of our research find applications in 3D digital effects and animation, the computer games industry and in computer-aided design and manufacturing (CAD/CAM).





Craig Gotsman

Distinguished Professor
craig.gotsman@njit.edu

Research Areas: Computer graphics, geometric modeling, geometry processing, computational geometry

Distance Computations on 3D Models and General Networks

Computing geodesic distances and paths on 3D manifolds is an important problem in geometry processing. Existing approximate algorithms may efficiently compute an entire “distance field” from a single point on the manifold, but computing arbitrary point-to-point geodesic distances efficiently is difficult. Our work takes advantage of an efficient (pre-processed) hierarchical subdivision of a discrete 3D mesh, when treated as a graph, based on compact and balanced cuts, to devise an efficient “divide and conquer” online algorithm to answer these queries very quickly. Our basic methodology is generalized in a number of ways: 1) To compute lower-bound heuristics for shortest-path distances on road networks, which enables efficient shortest-path computation with the A* algorithm. 2) To efficiently approximate resistance distances on social networks. 3) To perform compact message routing (using small routing tables) with modest stretch in sparse networks.



Przemyslaw Musialski

Associate Professor
przemyslaw.musialski@njit.edu

Research Areas: Computer graphics, geometric modeling, geometry processing, computational fabrication

Geometric Modeling and Computational Design

The traditional separation of design and manufacturing is currently one of the bottlenecks in the product development process leading to multiple expensive feedback loops. Our research aims at developing computational tools that integrate expert knowledge in 3D geometric modeling, physical simulation and smart algorithms in order to facilitate this process, which will become an important factor for success in increasingly competitive digital markets.

Computational Fabrication

We perform research at the intersection of applied mathematics, computer graphics and computational design. We investigate problems in geometric modeling and geometry processing, shape analysis and synthesis, as well as physical modeling and simulation with the goal of developing novel algorithmic solutions for the rapidly-growing domain of digital fabrication. Our research aims to provide novel solutions for the future design and manufacturing market.



Tomer Weiss

Assistant Professor
tweiss@njit.edu

Research Areas: Machine learning, data driven, optimization, simulation, crowd dynamics, deep learning, reinforcement learning

Computational Design of Virtual and Real Worlds

Virtual worlds are growing in complexity and interactivity and are challenging to construct efficiently and realistically. Interior scenes are an intrinsic part of virtual worlds. They are also needed for non-interactive applications, such as furniture marketing, residential showcasing and other commercial applications. Unfortunately, most of the aesthetic and creative process of designing interiors, real or virtual, is manual. To accelerate this process, we utilize machine learning and computer vision to develop several computational interior design tools. Our research assists users by providing suggestions on which furniture to pick, how well it matches the room’s current style and where to place it. Hence, we can facilitate content creation needs for multiple industries and, most importantly, accommodate our common human needs to create our own optimal interior environment.

Simulating the Motion of Multi-Agent Crowds

Crowd simulation is an essential part of multiple domains, from interactive media systems, to pedestrian analysis in evacuation scenarios and urban planning, to autonomous multi-agent path planning and coordination. A crowd is a collection of independent, self-actuated agents, where each agent has individual navigational goals in this shared environment. Agent movement is controlled by a navigation algorithm, which needs to ensure that an agent progresses towards its goal while avoiding collisions. Computing collision-free agent motion is difficult, due to the complexity of such dynamic interactions. Despite more than 30 years in crowd simulation research, simulation methods have subspecialized and are computationally effective for either sparse or dense crowds but not for both. Our work is the first method that allows real-time simulation of both dense and sparse crowds for up to 100,000 agents. This result is made possible by reframing agent motion as a constrained mathematical optimization problem that can be solved effectively on modern GPU computer hardware. Aside from the immediate practical implications for the gaming and visual effects industry, these results inspire future research on large-scale pedestrian dynamics which can affect urban and space planning.



Cybersecurity

The security research group performs extensive work in areas such as software supply chain security, web security and privacy, smartphone and mobile security, database security, trustworthy ML/AI and homomorphic encryption. It also makes significant contributions to open-source software. The group's research has been externally funded by agencies such as the NSF, DARPA, and the NSA, and members are recipients of two NSF CAREER awards and one DARPA Young Faculty Award. The group aims to publish its research in top security and related venues, such as Usenix Security Symposium, IEEE Security and Privacy Symposium, ACM Mobicom and ACM SIGMOD.





Reza Curtmola

Professor
reza.curtmola@njit.edu

Research Areas: Cybersecurity, software security, web security and privacy

Defending Software Supply Chains Against Hackers

Funded by DARPA and the NSF, we have developed in-toto, an open-source framework that promises to safeguard software for developers and end users. In-toto provides organizations with insights into the software development and distribution chain, such as having a provable assurance that proper software development practices were followed. With in-toto in place, it will be more difficult for malicious code to be slipped into software products, thus raising the bar significantly for attackers. Through integrations, in-toto is currently used by thousands of companies and has improved the security of millions of users.

Web Security and Privacy

The goal of this project is to explore targeted privacy attacks on the web through the lens of side channels. We uncovered new attacks that can lead to targeted deanonymization on the web by using CPU cache side channels. In particular, we uncover a set of practical and scalable attacks that can deanonymize users in several important settings for which prior attack methods are not effective. This affects all major browsers, including Chrome, Firefox, Safari, Edge, Tor Browser and numerous major sites, including Google, Twitter, LinkedIn, TikTok, Facebook, Instagram and Reddit. Our attacks run in less than 3 seconds in most cases and can be scaled to target an exponentially large number of users. More importantly, we provide a comprehensive countermeasure against all of the attacks we discovered. This countermeasure is already available on the Chrome and Firefox extension stores and can be downloaded and installed immediately by concerned users. This work was published in the 31st USENIX Security Symposium.



Iulian Neamtiu

Professor
iulian.neamtiu@njit.edu

Research Areas: Programming languages, software engineering and their applications to reliable AI, smartphones, security

Android Security

Our research is focused on security issues in Android apps and the Android platform, including exposing deceptive practices in apps, apps attempting to cover their traces, ransomware, unauthorized collection and transmission of user data in general and personally identifiable health information in particular, apps refusing to disclose the data they collect or refusing to delete data when legally mandated.



Kurt Rohloff

Associate Professor
kurt.rohloff@njit.edu

Research Areas: Encrypted computing, lattice encryption implementation, homomorphic encryption, cryptographic program optimization

Combating Data Leaks: PALISADE

We developed a widely used open-source lattice encryption library and software engineering tools for a new family of encryption technologies. This software library provides encrypted computing capabilities such as homomorphic encryption, allowing organizations to outsource computation to cloud computing environments without risking privacy and leaking sensitive information to potential adversaries.

Kurt Rohloff

Improving Usability of Open-Source Software

Funded by the first DARPA Young Faculty Award at NJIT and the IARPA HECTOR project, our MARSHAL and Verona projects focus on making it easier to rapidly optimize open-source software to run on commodity hardware. This research focuses on the deployment of the PALISADE open source lattice encryption library on embedded systems.



Shantanu Sharma

Assistant Professor
shantanu.sharma@njit.edu

Research Areas: Database, security, privacy, blockchain, IoT

Information-Theoretically Secure Processing

Despite over two decades of research, secure data outsourcing remains an open challenge. Information-theoretically secure techniques provide the highest level of security regardless of the computational capabilities of an adversary. One of the well-known information-theoretically secure techniques is Shamir's secret sharing. We develop information-theoretically secure data processing systems that can efficiently execute different types of SQL queries on large databases. Furthermore, we focus on information-theoretically secure machine learning techniques.

Smart and Privacy-Preserving Smart Spaces

Smart spaces are rapidly growing in present time. Examples of smart spaces are office/university buildings, shopping malls, train/bus stations and airports that capture user-related data via different types of sensors. While such sensor data is beneficial to developing multiple value-added services, smart spaces jeopardize user privacy due to mixing sensor data with the digital representation of space. For example, tracking a person in real-time can reveal their behavior. We develop an end-to-end secure and privacy-preserving smart space that respects user privacy at each stage of data processing, such as data collection, storage, processing, sharing and auditing.



Cong Shi

Assistant Professor
cong.shi@njit.edu

Research Areas: Mobile security, robust and trustworthy machine learning

Audio-Domain Position-Independent Backdoor Attack via Unnoticeable Triggers

Deep learning models have become key enablers of voice user interfaces. In this project, we aim to investigate an effective yet stealthy training-phase (backdoor) attack in the audio domain, where hidden/unnoticeable trigger patterns are injected through training set poisoning and overwrite the model's predictions in the inference phase. An attacker can simply play an unnoticeable audio trigger (e.g., bird chirps, foot steps) into live speech of a victim to launch the attack.



Data Mining

The abundance of data available, like images, natural languages, social networks and trajectories, has created opportunities for educators and governments to gain previously unavailable insights through a deeper understanding of the data and make better predictions with real-world datasets. Our group is focused on mining knowledge from data and developing tools that improve the process of data analytics. These analytics have created real-world, trustworthy intelligence. We work on a variety of techniques, including pattern mining, deep learning, generative models and reinforcement learning. These techniques provide tools to address issues ranging from data-driven simulation to fake news detection and time series data analysis.





Dimitri Theodoratos

Associate Professor
dimitri.theodoratos@njit.edu

Research Areas: Data mining, pattern extraction

Mining and Summarizing Patterns from Large Trees and Graphs

Extracting frequent patterns hidden in trees and graphs is critical for analyzing data and a first step for downstream data mining. Most pattern-mining algorithms do not scale to big data applications. We have designed algorithms to extract patterns from large trees and graphs, leveraging results using compressed bitmap views.

Searching Structured and Semi-Structured Data with Keyword Queries

Disambiguating a user's intention in posing a keyword query and efficiently retrieving relevant results is an immense challenge for keyword search when using big data. We have devised an approach that exploits a structural summary of the data to extract pattern graphs for keyword queries. This empowers non-expert users to extract information from data sources and databases without mastering a query language and without any knowledge of the organization or structure of data sources.



Jason Wang

Professor
jason.t.wang@njit.edu

Research Areas: Data mining, machine learning, deep learning, data science

Mining Big Data Through Deep Learning

We are designing and implementing new deep learning algorithms and architectures for mining big data. We have developed a 3D-atrous convolutional neural network, used it as a deep visual feature extractor and stacked convolutional long short-term memory networks on top of the feature extractor. This allows us to capture not only deep spatial information but also long-term temporal information in the data. In addition, we use stacked de-noising autoencoders to learn latent representations of the data that construct feature vectors suitable for classification. We also develop new recurrent neural networks to mine time-series data for stock market forecasting and space weather prediction. Currently, we are building a deep learning framework with generative adversarial networks. This framework will be used for stochastic video prediction, image synthesis and image-to-image translation. The framework can handle model uncertainty as well as data uncertainty and sparsity. Our deep learning models are suited for big data applications that have few, incomplete, imperfect, missing, noisy or uncertain training data.



Hua Wei

Assistant Professor
hua.wei@njit.edu

Research Areas: Data mining, reinforcement learning, urban computing

Learning Realistic Simulations from Real-World Data

This project aims to build a realistic traffic simulator by investigating data mining algorithms and provide solutions toward mimicking real-world simulations with real-world data. It focuses on the application context of learning to simulate the movements of humans, including human travelers or vehicles with human drivers, which essentially leads to a more realistic traffic simulator. The simulator of human movement is a starting point to build a city simulator. City simulators can be a valuable tool to quantify and optimize city policies, such as those for traffic signal control strategies. The city simulator can utilize multi-source urban data and advanced learning techniques to find better policies to make the city more sustainable.



Brook Wu

Associate Professor
yi-fang.wu@njit.edu

Research Areas: Text mining, information extraction, information retrieval

Early Detection of Fake News on Social Media

A major challenge of effective and early detection of fake news is fully utilizing the limited data observed at the early stage of news propagation. We propose a novel deep neural network to detect fake news early, by combining user and post-based features into status sensitive crowd responses. Experimental results show that our proposed model can detect fake news with greater than 90% accuracy within five minutes after it starts to spread and before it is retweeted 50 times. Most importantly, our approach requires only 10% labeled fake news samples to achieve this effectiveness under PU learning settings. We plan to extend this work by incorporating additional social context data extracted from user interactions to further enhance user representations and prediction accuracy.

Neural Fake News Detection

Motivated by the inevitability of 'neural' fake news, we are working on building a framework to generate indistinguishable neural fake news stories. Our ultimate goal is to use them to augment fake news training data to accurately detect neural fake news stories. In our design, the framework will have three components: **1. Synthetic News Generation** using a short claim as an input to neural language models. We are trying to resolve apparent contradictions and inconsistencies in the synthetic news generated by other approaches. **2. Deceptive Fake News Generation** by using fact tampering attacks on the generated news or fact tampering attacks on the claim. **3. Neural Fake News Detection.** Finally, we plan to use the generated fake news as training data, along with other publicly available true news, to train a neural fake news detection model.



Extended Reality

Our research covers the whole extended reality spectrum (e.g., virtual reality, augmented reality, mixed reality), gaming, 3D graphics, digital media and simulations. Our groups are as diverse and innovative as one would expect from researchers engaged in these rapidly evolving fields. We are focused on perception and cognition in simulated and/or augmented environments across different domains, such as healthcare, education, manufacturing and defensive military. We develop new and better technology and methodology for improved usability, decision-making and learning outcomes across multidisciplinary domains and diverse users. Our work supports innovation by utilizing psychophysiological and behavioral sensors such as facial emotion-sensing and eye tracking, as well as employing empirical user studies and information theory.





Frank Biocca

Professor
frank.biocca@njit.edu

Research Areas: Virtual and augmented reality systems, components for brain-computer interfaces, real-time public opinion measurement

Design of Virtual Environments and Interfaces to Support Information, Perception and Cognition

Our research examines the design of virtual and augmented reality hardware, software interaction techniques and applications to augment or change user thinking and cognitive performance. This research is done with teams within the distributed Media Interface and Network Design labs working within the context of medical, scientific and defensive military applications. For example, in an international project collaborating with teams in Spain and Korea, we examine how a brain-computer augmented reality environment can support and trace complex consumer decision-making, political decision-making and psycho-therapeutic applications. Users of the augmented reality environment also wear untethered brain sensors, other psychophysiological sensors and behavioral sensors such as facial emotion-sensing, to detect and measure how features of the virtual environment influence thinking. These neurocognitive indicators are used to adapt the virtual environment to the user in real-time. Other projects in this research stream involve virtual environments that change the perception of the user's body or how social cognition is affected by immersive, hyper-realistic avatars during negotiation, training and decision-making. The design of virtual environments to augment cognition also applies to interactive scientific visualizations. These augmented reality environments make physical forces or microscopic phenomena experienceable by the senses and allow the user to perceive relations in data. In a recent astrophysics project, we created an augmented reality environment that uses streaming data from head-worn, magnetic sensors to directly see and experience Earth's local magnetic fields.



Jacob Chakareski

Associate Professor
jacob.chakareski@njit.edu

Research Areas: Immersive communication, augmented/virtual reality

Virtual Human Teleportation

Virtual reality and 360-degree video are emerging technologies that can enable virtual human teleportation to any remote corner of the globe. This requires ultra-low latency, gigabit-per-second wireless speeds and data-intensive computing. Our research investigates synergies at the intersection of 6DOF 360-degree video representation methods, edge computing, UAV-IoT, millimeter-wave and free-space optics wireless technologies. It transmits data using much higher electromagnetic wave frequencies to enable the ultra-high data rates and ultra-low latencies required by next generation societal VR applications.

Real-Time Structure-Aware Reinforcement Learning

Reinforcement Learning (RL) provides a natural paradigm for decision-making in diverse emerging applications that operate in unknown environments and with limited data of unknown stochastic characteristics. Paramount to the effective operation of these ultralow latency applications, such as IoT sensing, autonomous navigation and mobile virtual and augmented reality, is the ability to learn the optimal operation actions online and as quickly as possible. Existing state-of-the-art RL methods either take too long to converge or are too complex to deploy. Our research examines novel structure-aware RL methods that integrate basic system knowledge to compute learning actions updates across multiple states or even the entire state-space of the problem of interest, in parallel. To address the challenge of computational complexity that is introduced, our methods integrate analysis that help effectively trade-off learning acceleration and computing complexity.

Jacob Chakareski Societal Applications

Our research focuses on interdisciplinary synergies to enable next-generation applications. For instance, a National Institutes of Health project at the intersection of networked virtual reality, artificial intelligence and low-vision rehabilitation aims to enable novel, previously inaccessible and unaffordable health care services to be delivered broadly and affordably. Other projects include the integration of virtual reality, real-time reinforcement learning and soft-exoskeletons for future physical therapy and the synergy of UAV-IoT and VR towards next generation forest fire monitoring.



Salam Daher

Assistant Professor
salam.daher@njit.edu

Research Areas: Augmented, virtual and mixed reality; physical-virtual, 3D graphics; virtual humans; synthetic reality; modeling simulation and training; distance simulation; healthcare applications and virtual patients

Mixed Reality Simulations to Improve Training

Our research focuses on creating simulations using computer graphics, multimedia and mixed reality to improve training in different domains including health care simulation. We are especially interested in research involving virtual humans and multisensory experiences. We developed a new class of augmented reality patient simulators called physical-virtual patients that allow health care educators to interact with a life-size simulated patient by providing real-time physical tactile cues such as temperature and pulse; auditory cues such as speech and heart sounds; rich dynamic visual cues such as facial expressions indicating pain or emotions; and changes in appearance such as skin color and wounds.

Training Caregivers of Virtual Geriatric Patients

We are developing a new generation of Virtual Geriatric Patients (VGP). The VGPs are realistic, embodied, conversational virtual humans who are aware of their surroundings. The VGPs are displayed in Mixed Reality as training scenarios aimed to improve caregivers' perceptions, attitudes, communication, and care towards older adults. This research is supported by a grant from the National Science Foundation Future of Work at the Human Technology Frontier.

3D Graphics for Wound Visualization, Measurements and Tracking

Our research focuses on visualizing wounds in 3D for accurate measurements, reduced variability of measurements and improved tracking of patients' progress. In the clinical setting, this translational research can reduce errors, improve healing estimates, and improving patient outcomes. In the training setting, this technology can improve healthcare trainees' skills in wound assessment, especially when combined with mixed reality. This research is partially supported by the New Jersey Health Foundation.

Interactive Remote Simulation for Healthcare Training

During the pandemic, healthcare educators rushed to use pre-existing videos or had to record their own videos that they shared with their students to watch as a makeshift "simulation". Content needs to be interactive to satisfy the interactivity requirement for simulation. Our team developed software called "Anywhere Simulation (AwSIM)" that allows healthcare educators to add interactivity to their existing content to create new healthcare scenarios and share that content remotely with their students. AwSIM provides healthcare educators with the power to create their own simulation scenarios using their own content (e.g., videos, images, text) and run the simulation remotely with their students. The software is content independent and is easy to use. We ran multiple studies with nursing students using our AwSIM technology and found that adding interactivity promotes teamwork, perception of authenticity and higher levels of thinking. Also, the AwSIM software has a high technology acceptance rate among

Salam Daher

students. We are working on creating and evaluating an immersive standalone version for trainees that they can use on a flat screen or on with a head mounted display with or without a facilitator.



Margarita Vinnikov

Assistant Professor
margarita.vinnikov@njit.edu

Research Areas: Immersive and collaborative cross reality, navigation, gaze/body tracking

Immersive Cross-Reality Applications

Our research examines general areas of interactive cross-reality (XR), which encompasses virtual, augmented and mixed reality applications and serious game development. We specialize in eye and body tracking as well as multi-sensory augmentations. Specific topics include the design, development and evaluation of novel XR and cross-model (visual, audio and/or haptic) user experiences through simulations such as walking in European cities and driving in New Jersey and New York. We also build augmented-reality collaborative applications.

Multi-User Gaming and Collaborative Platforms

Virtual collaboration received a lot of attention recently as many people are forced to work away from their usual workspaces due to the COVID-19 pandemic. Providing a realistic environment where people can reliably and efficiently collaborate on tangible objects and models will help many businesses. Primarily, this is relevant to city planners, military and law enforcement, as well as educational settings. We are also interested in a multi-calibration platform between various augmented and mixed reality devices such as mobile phones, Hololens and Magic Leap.

Visualization of Large Datasets in Virtual and Augmented Reality

Large data sets such as ontology trees or visibility graphs, when loaded into virtual or augmented reality devices, can pose many challenges. For example, continuous loading of data into a mobile device. Similarly, there are no established methods for the most userfriendly way to visualize large data clouds. Hence, we combine various computer science algorithms with user-studies to develop the most efficient ways to visualize large data sets.

A photograph of two women in a server room. The woman on the left, with dark curly hair, is looking at a tablet held by the woman on the right, who has long blonde hair. They are both wearing white shirts and blue lanyards. The background shows server racks and glass partitions.

Human-Computer Interaction

Human-computer interaction (HCI) is a multidisciplinary field of study that has transformed how we interact with technology. The work draws on our diverse backgrounds in psychology, cognitive science, education, journalism, computer science, artificial intelligence and music. Our research contributes to topics and areas in HCI such as computer-supported collaborative work, social media, online communities, crowdsourcing and social computing, human-AI interaction, interactive audio, extended reality, information visualization, visual analytics, inclusive technology, educational technology and accessibility. This work is frequently published in top-tier venues such as ACM CHI, ACM CSCW, ACM ASSETS, TVCG and IEEE VIS.



Julie Ancis

Distinguished Professor
julie.r.ancis@njit.edu

Research Areas: Cyberpsychology, human-computer interaction, tailored social media messaging, disinformation, bias

Social and Behavioral Impacts of Online Information

The informational landscape and how individuals and groups engage with this information has implications for a broad range of human behaviors. Using qualitative and quantitative approaches grounded in psychology and computing, we analyze how both the text and visual components of social media messages are constructed to convey information about issues ranging from health crises to critical societal events. Our research investigates the impact of these messages on risk perception, decision making, proactive health behaviors, civil discourse, and bias. This work has four major foci: **1.** The evaluation of tailored and targeted online messaging. **2.** Its impact on psychological processes and human behavior. **3.** Development of culturally informed online information and messages to enable adaptive behavior. **4.** Testing the psychological and behavioral impact of culturally relevant messaging on perceptions and intentions.



Mark Cartwright

Assistant Professor
mark.cartwright@njit.edu

Research Areas: Machine listening, interactive machine learning, human-computer interaction, audio processing, music information retrieval

Tools for Accessible Sound Understanding and Creation

Sound is all around us and in the media we consume. However, for the roughly 1 in 5 people who are to some degree deaf or hard of hearing (DHH), the information contained in this medium is not always accessible. In this project, we aim both to develop tools to increase the accessibility of information contained in sound and to develop tools to make the creation of audio artifacts more accessible.



Aritra Dasgupta

Assistant Professor
aritra.dasgupta@njit.edu

Research Areas: Information visualization, human-centered data science, science communication

NJIT's Intelligible Information Visualization Lab (NiiV)

As George Bernard Shaw so eloquently said: "The single biggest problem in communication is the illusion that it has taken place." This is so evident in today's age, where information, if communicated properly, can cure diseases and fuel discoveries, but, if miscommunicated, can lead to an "infodemic" in the worst-case scenario. To solve this conundrum, at NiiV, we pursue intelligibility as the foundational principle for making information more accessible, meaningful and actionable to domain experts (e.g., doctors, climate scientists, policy-makers, etc.) and non-experts alike. We operationalize this principle by visualizing data, big or small, with the ultimate goal of letting human observers see, understand and trust the information that is often generated by black-box algorithms. By embracing a human-centered data science approach that ultimately culminates in interactive visual analytic interfaces, we preserve the best of both worlds: the power of computational methods and that of human judgment and reasoning. Our research has been supported by grants from the National Science Foundation (NSF) and the U.S. Department of Energy.



Amy Hoover

Assistant Professor
amy.k.hoover@njit.edu

Research Area: Optimization

Methods in Quality Diversity

While many optimization techniques in evolutionary computation maximize fitness with respect to one or more objectives, often such searches restrict the set of candidate solutions to those with objective values lying along the Pareto front of optimality. Instead, algorithms in Quality Diversity (QD) fully explore these objectives by specifying them explicitly as dimensions (called behavioral characteristics) that are characterized by their genomic, phenotypic, or behavioral traits. Such exploration can not only generate a large collection of high-performing solutions, but with well-chosen dimensions can potentially find higher performing solutions than pure-objective based searches alone.



Sooyeon Lee

Assistant Professor
sooyeon.lee@njit.edu

Research Areas: Human-computer interaction, accessibility, human-AI interaction

Design and Evaluation of Accessible AI Technologies for Users with Disabilities

Over one billion people in the world live with some type of disability. Many of them experience barriers in accessing information or using technologies, which can limit social interactions in both physical and digital spaces. Our work investigates the diversity of users, explores and leverages emerging technologies and adopts human-centered AI and inclusive design approaches in the design and evaluation of new AI based systems and applications that address accessibility barriers. Our research focuses on investigating and designing non-visual interaction for the community of blind users and non-audio and non-speech interaction for the community of deaf and hard of hearing users.



Yvette Wohn

Associate Professor
wohn@njit.edu

Research Areas: Human-computer interaction, social media, digital economy

Social Media Content Moderation

You may think the Internet is filled with bad content, but things would be much worse if not for the invisible work of armies of people who try to keep the Internet a pleasant space. Supported by the National Science Foundation, this work focuses on the work of content moderators on various social media platforms and understanding how to develop and maintain safe spaces online.

New Digital Economies

The metaverse may seem like a new phenomenon but research on virtual spaces has been going on for decades. In this research, we focus on systems with novel digital economies in virtual environments that have unique digital currency, such as online games and digital patronage systems. Our research examines spending behaviors and exchange patterns of virtual goods in games and other alternative financial platforms and how these activities are tied with creative content generation.

Fandom and Online Communities

This project examines how fans around the world use social media to engage in various collective action, including charity projects and political activism.

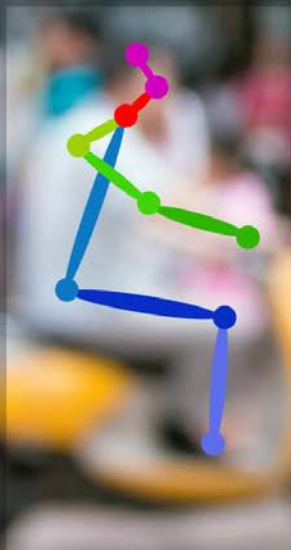


Image and Video

Our research on image and video processing continues to advance science and technology. We work with experts in health care, state and local government, security and the private sector to examine obstacles and address them by leveraging the latest advancements. From the diagnosis of pneumonia and brain tumors to automated systems that can monitor traffic incidents, roadway congestion and public safety risks, this group is making advances with significant work in image forensics and watermarking, pattern and face recognition, medical image analysis, video analytics, deep neural networks and smart transportation systems.



Chen Liu

Professor
chengjun.liu@njit.edu

Research Areas: Computer vision, pattern and face recognition, video processing

Making Intelligent Transportation Systems Smarter

The New Jersey Department of Transportation designated more than 400 closed-circuit video cameras statewide for incident monitoring, traffic congestion control and public safety operations. Video streams from these cameras feed to a back-end system. There, video analytics software is used to perform target detection and incident monitoring applications. We are actively working on incorporating wireless sensor networks, hierarchical edge-computing and computer vision to mitigate the challenging problems in various illumination and weather conditions in order to achieve fast and automated video based traffic monitoring.

Video Analytics Pilot Studies and Testing of Technologies

We propose a new modular approach for statistical modeling of traffic incidents and model selection in order to improve state-of-the-art traffic detection and monitoring. We investigated and developed automated video analytics systems to replace human operators for traffic incident detection and to monitor the cameras installed along the major New Jersey highways. We test the proposed technologies and benchmark their performance.



Frank Shih

Professor
frank.y.shih@njit.edu

Research Areas: Image processing, artificial intelligence, digital watermarking, digital forensics, robot sensing, neural networks

Deep Learning for Pneumonia Classification and Segmentation on Medical Images

Automatic identification of pneumonia on medical images has attracted intensive study. In this project, we develop a novel joint-task architecture that can learn pneumonia classification and segmentation simultaneously. Two modules, including an image preprocessing module and an attention module, are developed to improve both classification and segmentation accuracies. Experimental results performed on the massive dataset of the Radiology Society of North America have confirmed its superiority over other existing methods.

Deep Morphological Neural Networks and Applications

Given a target image, determining suitable morphological operations and structuring elements is a cumbersome and time-consuming task. In this project, we propose new morphological neural networks, which includes a nonlinear feature extraction layer to learn the structuring element correctly, and an adaptive layer to automatically select appropriate morphological operations. We also use them for their classification applications, including hand-written digits, geometric shapes, traffic signs and brain tumors. Experimental results show higher computational efficiency and higher accuracy when compared with existing convolutional neural network models.

Machine Learning, Deep Learning and Reinforcement Learning

Our research areas cover a wide range of topics in machine learning and deep learning, including reinforcement learning, graphical neural networks, adversarial learning, interpretable machine learning and natural language processing. Our application fields include FinTech, Blockchain, transportation, urban computing, bioinformatics and medical informatics, social network analysis and systems. Many machine learning faculty members are also core faculty of the Center for AI Research (CAIR). Our research has been supported widely by different agencies, including NSF, DOT, NIH, DOD and industry. The results are continually published in high-impact journals and prestigious machine learning and artificial intelligence conferences such as Nature Machine Intelligence, Nature Communications, IEEE Transactions on Neural Networks and Learning Systems (TNNLS), NeurIPS, ICML, KDD, AAAI, IJCAI and ICDM.





Jing Li

Assistant Professor
jingli@njit.edu

Research Areas: Real-time systems, parallel computing, cyber-physical systems and reinforcement learning for system design and optimization

Reinforcement Learning-Based System Design

The design space of modern complex systems is increasingly large. Finding good designs often involves solving mixed-integer optimization problems that are highly intractable. Our research develops reinforcement learning-based frameworks with graph neural networks and active learning techniques to intelligently and efficiently find good designs from the huge design space. We have applied our frameworks to various system design tasks, including resource allocation in cyber-physical systems and circuit design.



Yao Ma

Assistant Professor
yao.ma@njit.edu

Research Areas: Machine learning and data mining

Towards Advanced Filtering and Pooling Operations for Graph Neural Networks

In recent years, we have witnessed a rapid growth in our ability to generate and gather data from numerous platforms in the online world and various sensors in the physical world. Graphs provide a universal representation for a variety of data including online social networks, knowledge graphs, transportation networks and chemical compounds. Entities can usually be represented as nodes while their relations can be denoted represented as edges. Many important real-world applications on these data can be treated as computational tasks on graphs. A crucial step to facilitate these tasks is to learn good vector representations either for nodes or graphs. Recently, graph neural networks, which generalize deep learning techniques to graphs, have been widely adopted to learning representations for graphs. Though graph neural networks have advanced numerous real-world applications from various fields, they still suffer from many limitations in terms of efficacy and efficiency. This project aims to address these limitations by conducting theoretical analysis and developing innovative algorithms. This project is specifically motivated by applications to computational social science, computational biology and fraud detection in e-commerce.



Usman Roshan

Associate Professor
usman.w.roshan@njit.edu

Research Areas: Machine learning, medical informatics

Adversarial Robust Machine Learning With 0-1 Loss:

Machine learning models today are highly accurate but not very robust. They can be fooled to misclassify data with minor perturbations known as adversarial attacks. Adversaries targeting a given convex model are known to affect other convex models. We find this transferability phenomenon to be less effective between 0-1 loss and convex losses such as hinge and logistic, both of which are approximations to 0-1 loss and known to be affected by outliers. Consequentially, it is harder to attack 0-1 loss models with convex substitute model black-box attacks, and when the black-box attacker is 0-1 loss, the attack is highly ineffective on all models. Based upon these observations, we are researching novel algorithms and design implementations for scalable and faster 0-1 loss models.



Hai Phan

Assistant Professor
hai.phan@njit.edu

Research Areas: Social network analysis, machine learning, spatio-temporal data mining

Ontology-Based Interpretable Deep Learning

Machine learning models are trained with large amounts of data and achieve a certain level of competency in interpreting and classifying new input data. However, even if they work well, it can be difficult to explain why. Linger doubt persists that, in some situations, the classification output of the model might be wrong. In applications such as self-driving cars, this could have spectacularly negative consequences. We tie predictions of the model to a set of keywords taken from a predefined vocabulary of relevant terms. The number of words hard-coded into the model that influence the outcome produced by a machine learning for a new input is reduced and those words are taken from a limited and relevant ontology. This makes the output of the model easier to interpret as it becomes independent from terms that are irrelevant to the application domain.



Guiling "Grace" Wang

Distinguished Professor
gwang@njit.edu

Research Areas: Deep learning, FinTech, blockchain technologies, intelligent transportation systems

Deep Reinforcement Learning for Intersection Control

Reinforcement Learning-based traffic signal controllers can adaptively adjust signals based on real-time demand. To learn a good policy, we have proposed combining a dueling network, target network and double Q-learning network with prioritized experience replay technology in one signalized intersection. It has proven to be a successful attempt to stabilize the learning process and mitigate over-estimations of the learning agent. To further guarantee the safety of vehicles, we have incorporated domain safety standards in the above RL-based traffic signal controller. Nonetheless, the small proportion of the collision data makes this problem extremely challenging, not to mention that the learning agent might not obey the safety rules in practice. We thus, instead of letting the RL model learn by itself, incorporate domain safety standards into different parts of the RL model (i.e., action, state, loss function and reward function) as a safety shield. This safety-enhanced approach can guide and correct the RL agent toward learning a much safer action and dramatically lower the collision rate. Meanwhile, we not only focus on one intersection but also consider multiple intersections as a network to make a smoother driving experience. By enabling cooperation among intersections, each RL agent learns to communicate with others to spread out the traffic pressure quickly. Intensive experiments show the effectiveness of our systems.



Lijing Wang

Assistant Professor
lijing.wang@njit.edu

Research Areas: Machine learning, deep learning, natural language processing

Fine-Tuning for Improved Bias and Variance for Information Extraction

The bias-variance tradeoff is the idea that learning methods need to balance model complexity with data size to minimize both under-fitting and over-fitting. Recent empirical work and theoretical analysis with over-parameterized neural networks challenges the classic bias-variance trade-off notion. We are exploring a variance decomposition-based justification criteria to examine whether large pretrained neural models (e.g., RoBERTa) in a fine-tuning task are generalizable enough to have low bias and variance.

Machine, Deep and Reinforcement Learning



Hua Wei

Assistant Professor
hua.wei@njit.edu

Research Areas: Data mining, reinforcement learning, urban computing

Reinforcement Learning in Urban Environments

Reinforcement learning (RL) is a general decision-making paradigm and has recently made steady progress in academia. We are tackling a series of unique challenges for real-world systems like urban environments, including incomplete and noisy data, trustworthy decisions making, and discrepancies between the learning environment and reality.

Securing Multi-Agent Reinforcement Learning with Adversarial Attacks and Graph Learning

Multi-agent reinforcement learning (MARL), the extension of reinforcement learning (RL) methods for multi-agent domains, has gained popularity for generating high-quality solutions in some domains, such as autonomous racing vehicles and multi-robot control. As its popularity increases, so do the security issues arising from adversarial attacks, where an adversary can cause a well-trained agent to behave abnormally by tampering with the input to the agent's policy network or training an adversarial agent to exploit the victim's weakness. This project will develop a number of novel methods to ensure the robustness of RL by exploring adversarial attacks and training antagonistic agents for competitive MARL.



Zhi Wei

Professor
zhiwei@njit.edu

Research Areas: Machine learning, statistical modeling, bioinformatics

Explainable AI for Unsupervised Learning

In recent years, explanation techniques have emerged and received a lot of attention. In supervised learning settings, it emphasizes the ability to correctly interpret a prediction model's output. Most existing explanation works have been focused on supervised learning, such as LIME and SHAP. Yet very little work has been done in an unsupervised learning setting. Our goal is to explain why and how a sample is assigned to a specific cluster in unsupervised learning. Specifically, we would like to know which features are portrayed as contributing to a cluster, or evidence against it. With this information, a practitioner can make an informed decision about whether to trust the model's cluster assignment. There is also a so-called "double use of data" problem when trying to find discriminative features that distinguish the resultant clusters. We will apply the new methods to analyze finance and accounting data.





Mobile Computing

Mobile computing and sensing have become ubiquitous in the world, due to the widespread availability of smart phones, smart watches and IoT devices. The availability of computing and sensing anytime, anywhere presents opportunities for novel applications that can improve our daily life, in domains such as mobile health, transportation, communication and safety. We focus on applying machine learning techniques to solve mobile computing and sensing challenges. The results of this research have been published in top venues such as ACM SenSys, AAAI and IEEE Transactions on Mobile Computing.



Cristian Borcea

Professor
borcea@njit.edu

Research Areas: Mobile computing and sensing, machine learning for mobility, predictive computational advertising

Federated Learning for Mobile Sensing Data

Federated Learning (FL) is a distributed machine learning paradigm that enables privacy-aware training and inference on mobile devices with help from the cloud. FL can enable a wide range of mobile apps that use machine learning models on mobile sensing data. We created FLSys, a mobile-cloud FL system that can be a key component of an ecosystem of FL models and apps. FLSys balances model performance with resource consumption, tolerates communication failures, achieves scalability, provides advanced privacy mechanisms and supports third-party apps and models. We also created a system for fine-grained location prediction of mobile users, based on GPS traces collected on smart phones. Applications that benefit from this system include video quality adaptation as a function of the 5G network quality at the predicted user locations and augmented reality apps that speed up content rendering based on predicted user locations.



Cong Shi

Assistant Professor
cong.shi@njit.edu

Research Areas: Mobile security, robust and trustworthy machine learning

Solving the WiFi Sensing Dilemma in Reality Leveraging Conformal Prediction

WiFi sensing has demonstrated its great convenience and contactless sensing capabilities in supporting a broad array of applications. However, designing a ubiquitous WiFi sensing system for heterogeneous scenarios in practice is still a big dilemma as the system performs poorly under domain variations. In this project, we aim to investigate reliable WiFi sensing based on a statistical assessment approach, named conformal prediction. The proposed approach quantifies the conformity between new testing WiFi samples and the training samples for prediction, which enhances the reliability of deep learning models without generating new features or retraining.



Software Engineering and Programming Languages

Software pervades all aspects of modern life; and so do software failures. Whereas in other engineering disciplines product correctness and reliability are the norm, in software engineering software failures are a fact of life. The price tag of this inadequacy is massive: A recent study by the Consortium for Information and Software Quality estimates that, as of 2020, the cost of poor quality software in the USA exceeds \$2.3 trillion, including \$1.5 trillion due to operational software failures and \$0.5 trillion due to poor quality legacy systems. Like all complex, multi-dimensional problems, the challenge of software quality cannot be addressed by any single approach. Our software engineering research includes a rich, diverse portfolio of approaches (analytical, empirical), paradigms (static, dynamic), quality attributes (correctness, reliability, security), lifecycle phases (testing, verification, repair) and products (methods, tools, artifacts).



Ali Mili

Professor
mili@njit.edu

Research Areas: Software engineering, software testing, program correctness, program repair

Functional Extraction

Despite several decades of research and technology transfer, the development of reliable and secure software remains an unfulfilled goal: software is routinely fielded with known failures but unknown faults. Many of the problems of software engineering stem from the fact that programmers develop and reuse software products whose function they do not fully comprehend. The purpose of the functional extraction project is to develop and evolve an automated tool that computes the function of a program by static analysis of its source code in a C-like language. We achieve this goal by mapping a program onto an equation between its inputs and outputs, which we can then solve or reason about to analyze correctness, safety, security, reliability, etc. The main innovation of this project is the ability to compute the function of loops, which we achieve by means of invariant relations.

Theoretical Foundations for Program Repair

To repair a program does not mean to make it correct; it only means to make it more-correct, in some sense, than it was. In this project we introduce a definition of relative correctness, a reflexive transitive partial ordering between candidate programs that culminates in absolute correctness with respect to a reference specification, and we show how this concept can be used in program repair to: define what is a fault; define what is a unitary/atomic fault; define what it means to repair a fault; measure the degree of faultiness of a program; and perform program repair in a stepwise manner, by removing one unitary fault at a time until the program is absolutely correct. We show empirically that by augmenting existing program repair tools with our capabilities, we enhance their performance and their scope.

Mutation Testing

Mutation testing is the art of generating mutants of a base program and is used primarily to assess the quality of test suites: a test suite is all the better that it can distinguish the base program from its non-equivalent mutants. We estimate the ratio of equivalent mutants of a base program by analyzing the amount of redundancy in the program. Also, we analyze the effectiveness of a mutant set by its ability to vet test suites that expose the failures of the base program; and we model the minimization of a test suite as an optimization problem, where the objective function is the cardinality of the set and the constraint under which the optimization is sought is the preservation of the mutant set effectiveness.



Martin Kellogg

Assistant Professor
martin.kellogg@njit.edu

Research Areas: Verification, automated reasoning, software engineering, programming languages

Inferring Type Qualifiers

Pluggable type systems enable programmers to enhance the ability of a programming language's type system to automatically prevent bugs by attaching a type qualifier to each type in the program: each pluggable type system prevents one kind of bug. These techniques have been shown to be effective, but can be difficult to apply to legacy systems because of the need to write type qualifiers. We have developed a novel type inference system that is general, meaning that it applies to any pluggable type system without needing to modify the pluggable type system itself, making the techniques more accessible to programmers.


```
< group_info
int len = cpcount * sizeof(*groupinfo);
int cpcount = min(NGROUPSPERBLOCK, count);
return groupinfo;
```

Martin Kellogg

Correlating Understandability and Verifiability

Intuitively, code that is simple for a human to understand ought to be simple to automatically verify as free of bugs. We have developed an experimental methodology to test for this correlation in real code and applied it to a corpus of programs used in prior studies of how humans understand code and found evidence for this intuition. Our findings have implications for how we build and use verification tools: when automated tools fail to verify code, it does make sense to refactor that code to make it simpler, because code that is simpler for the verifier to understand is more likely to be simple for other programmers to understand as well.



Iulian Neamtiiu

Professor
iulian.neamtiiu@njit.edu

Research Areas: Programming languages, software engineering and their applications to reliable AI, smartphones, security

Improving Smartphone Reliability and Security

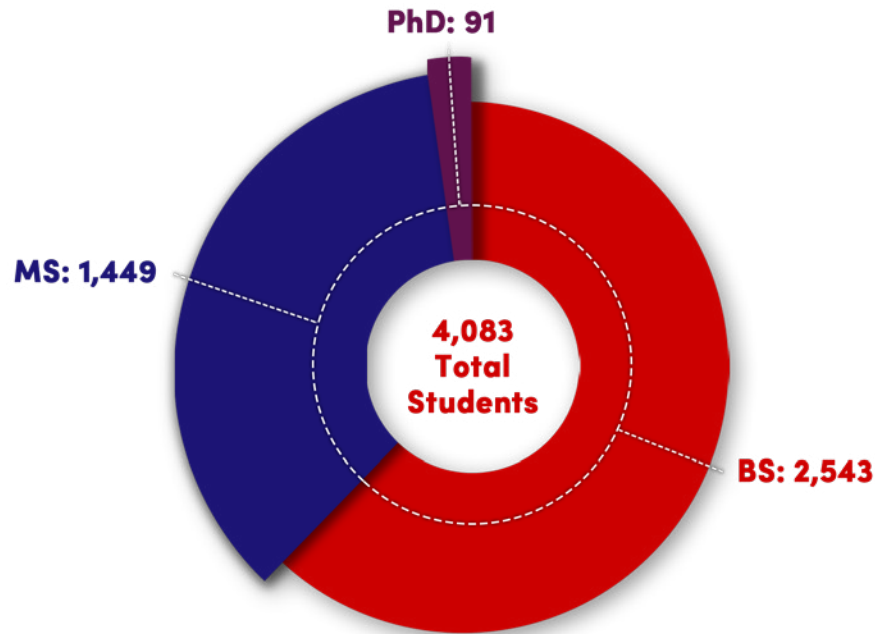
We have developed a wide range of approaches to improve smartphone reliability, including static and dynamic program analyses, record-and-replay systems, application self-healing and automatic test generators. These approaches were released as open-source code and have exposed reliability issues in many popular applications including medical mobile apps used in acute care settings.

Verifying Unsupervised Learning Implementations

Unsupervised Learning (UL) is widely used, by itself or as part of AI pipelines, in scenarios ranging from everyday applications to high-stake domains. However, even specifying UL correctness is a challenge, which complicates verifying or validating UL implementations. Our past work has exposed serious issues (e.g., nondeterminism, inconsistency) in widely popular UL toolkits and on critical, e.g., medical datasets. Our current work is focused on automated verification and fault localization for UL implementations using static and dynamic analysis.



COLLEGE OF COMPUTING FACTS AND FIGURES 2023



Departments

DATA SCIENCE

COMPUTER SCIENCE

INFORMATICS

Total Research Funding
2022-2023

\$8.5M

Tenure-Track Faculty

57

Lecturers

47

35%

of total NJIT Student Body

23%

Female Students



Undergraduate Degrees

B.S. Business and Information Systems
 B.S. Computer Science
 B.S. Computing & Business
 B.S. Data Science
 B.S. Human-Computer Interaction
 B.A. Information Systems
 B.S. Information Technology
 B.S. Web and Information Systems



Graduate Degrees

M.S. Artificial Intelligence (AI)
 M.S. Bioinformatics
 M.S. Business & Information Systems
 M.S. Computing & Business
 M.S. Computer Science
 M.S. Cybersecurity and Privacy
 M.S. Data Science
 M.S. Information Systems
 M.S. IT Administration & Security
 M.S. Software Engineering



Doctoral Degrees

Ph.D. Computer Science
 Ph.D. Data Science
 Ph.D. Information Systems



NJIT

COMPUTING

NJIT

HONORS

NJIT

MANAGEMENT

NJIT

YING WU COLLEGE
OF COMPUTING





YING WU COLLEGE
OF COMPUTING

Ying Wu College of Computing
New Jersey Institute of Technology
University Heights
Newark, NJ 07102

973-596-2822 | computing.njit.edu

