NJIT New Jersey Institute of Technology

YING WU COLLEGE OF COMPUTING RESEARCH REPORT

2018-2019

TABLE OF CONTENTS

NJIT
New Jersey Institute
of Technology

The Ying Wu College of Computing (YWCC) was established in 2001 and has grown rapidly to become the second-largest college at New Jersey Institute of Technology (NJIT), consisting of the Computer Science and Informatics departments. With close to 3,000 students at all levels and approximately 800 graduates per 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 in the last decade. These talented individuals are responsible for the high-quality academic research described in this report, research that is financially supported by a variety of government funding 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 the YWCC researchers ranges from deep mathematical theory to very 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,

C Gotsman

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

BIG DATA

Faculty at NJIT work on a variety of Big Data topics including dimension reduction, complex workflows and energy optimization, 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 researchers have presented their findings at top-tier conferences including VLDB and SIGMOD, and have received various awards including the SIGMOD Test of Time Award in 2015. Researchers have also received best paper awards at conferences for high-performance computing, parallel and distributed computing, multimedia and other relevant themes.

Vincent Oria

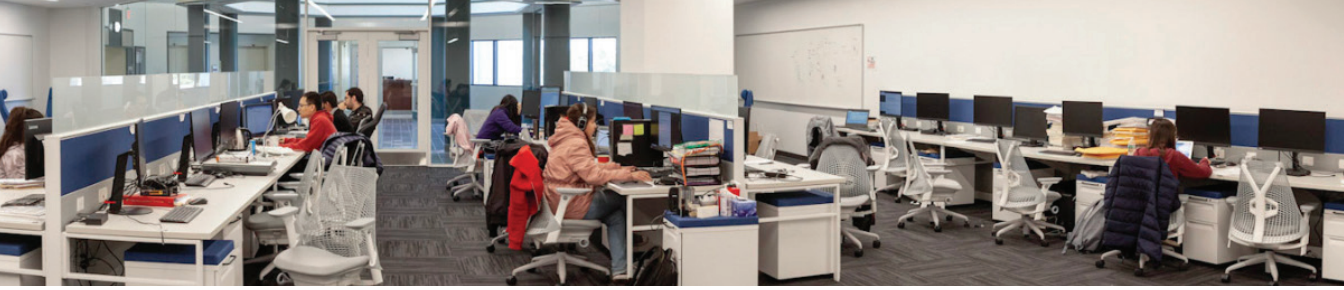
SEARCH AND CLUSTERING IN HIGH DIMENSIONAL SPACES

In many applications such as multimedia and recommender systems, data is often represented as vectors in high dimensional spaces, each dimension representing a feature of the data. When the number of features (the data dimensionality) is high, the ability to differentiate similar measures diminishes to the point where methods (such as search and clustering) lose their effectiveness. In collaboration with the National Institute of Informatics in Japan, we are investigating effective search and clustering methods for high-dimensional data that integrate novel feature selection techniques.

MANAGING SOLAR FLARE DATA

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 and enrich solar data captured by various solar flare observing instruments around the world.

VINCENT ORIA
Professor
Multimedia databases, spatio-temporal databases, recommender systems.
vincent.oria@njit.edu



Dimitri Theodoratos

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 innovative 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.

DIMITRI THEODORATOS
Associate Professor

Data mining and pattern extraction.
dth@njit.edu

CHASE WU
Associate Professor

Big data, green computing and networking, parallel and distributed computing.
chase.q.wu@njit.edu

Chase Wu

REVOLUTIONIZING PROCESSING OF BIG DATA SETS

To reduce processing time of information in complex scientific computing infrastructures, we develop practical computing and networking toolkits to improve efficiency of complex workflows in big data and high-performance environments.

Visual inspection can often detect complex models or discover new patterns in big data environments. We develop visualization algorithms for 3D-volume data generated by scientific computations on supercomputers. Visual feedback is critical to the understanding and validation of physics models used for simulations in computational sciences.

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 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. We develop reliable tools to detect and contain radioactive materials to protect the populace and reduce the risk of radiological dispersal devices, such as "dirty bombs."

DATA SCIENCE

The amazing abundance of data available today has created opportunities for corporations, educators and governments to gain previously unavailable insights through deep analysis of data sets. This knowledge discovery has created real-world, actionable intelligence.

Our faculty work on all aspects of data science including natural language parsing, machine learning and deep learning. These techniques provide tools to address issues ranging from security, to fake news detection, to healthcare and genomic data analysis. The work of our researchers has been published in a wide variety of venues including Nature, Science, Nature Medicine, Cancer Discovery, and Nature Communications, as well as top-tier data mining, machine learning, biostatistics, and bioinformatics conferences and journals.

Jason Wang

MINING BIG DATA THROUGH DEEP LEARNING AND FEATURE ENGINEERING

We are designing and implementing new architectures and algorithms for deep learning and feature engineering. We have developed a new 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 long-term dependencies in solar filament eruption videos. We are also developing new algorithms to infer the interactions of the features including physical, temporal, spatial and image features of solar events. Such interactions, called "feature networks," will help perform dimension reduction and feature selection. This overall approach can capture not only deep spatial information but also long-term temporal information in the videos.

USING DEEP LEARNING FOR SECURITY VULNERABILITIES

Our research includes detection of malware and security vulnerabilities using deep learning. We represent malware detection programs using function call graphs and develop a novel graph embedding technique, using unordered tree matching and bipartite matching that maps function call graphs to feature vectors. This new graph embedding technique is comparable to well-known techniques such as DeepWalk, node2vec and struc2vec while working effectively for malware and vulnerability detection.

JASON WANG
Professor
Data mining, database systems, bioinformatics.
wangj@njit.edu

Hai Phan

PRESERVING SENSITIVE HEALTHCARE 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 health care providers to legal jeopardy. We have created "DeepPrivate," a system that uses machine learning techniques to protect personal medical information against cyberattacks.

ENABLE REAL-TIME DRUG ABUSE RISK BEHAVIOR DETECTION

Some Twitter users are surprisingly willing to divulge private details of their lives, including their own drug abuse activities. However, such tweets are hidden within hundreds of millions of other posts, literally needles in a very large haystack. We apply machine-learning methods to recognize drug-abuse-related tweets in a large collection of Twitter postings. Many tweets are geographically tagged, and all are temporally labeled. This allows us to recognize hot spots of drug activities and peak days and times to provide near-real time information to public health officials to plan appropriate tactical and strategic responses.

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 very well, it is nearly impossible to say "why they work well" and lingering doubt persists that in a particular situation, the classification output of the model might be wrong. In applications such as self-driving cars, this could have spectacularly negative consequences. In our research, we tie predictions of the model to a set of keywords taken from a predefined vocabulary of relevant terms. In other words, the number of words "hardcoded into the model" that influence the outcome produced by a machine learning model for a new input is reduced, and those words are taken from a limited and relevant collection of terms (an ontology). This makes the output of the model easier to interpret as it becomes independent from terms that are irrelevant to the application domain.

HAI PHAN
Assistant Professor
Social network analysis, machine learning, spatio-temporal data mining.
hai.phan@njit.edu

DATA SCIENCE

Usman Roshan

MACHINE LEARNING FOR GENOMICS AND MEDICINE

Our research focuses on developing new machine learning algorithms for cancer risk prediction from whole genome data for classification and alignment of DNA and protein sequences. This has broad applications leading to a better understanding of biology and accurate medical diagnoses.

RANDOM WEIGHTS AND WEAK TRAINING IN CONVOLUTIONAL AND STANDARD NEURAL NETWORKS

We are exploring new strategies to create faster and simpler training methods for neural networks, which are widely used today for classification tasks. Starting with random and weakly trained weights, as opposed to having all weights initially equal, we create architectures with random weights that attain benchmark accuracies comparable to leading neural networks and are less sensitive to gradient-based adversarial attacks than traditional networks.

Zhi Wei

ADVANCED ANALYTICS AND LEARNING FOR GENOMICS

We are developing advanced data analytics, data mining techniques, and machine learning modeling, with applications to a broad range of fields including biology, medicine and healthcare, digital marketing, and finance. We are especially concerned with how model-driven approaches, theories, or empirical applications can be used to address various challenges arising from exploiting large data sets. Current research topics of interest include, but are not limited to, probabilistic and statistical models and theories, deep learning, machine learning and data mining theories, models and systems. A focus area is analysis of genomic and genetic data, with application for cancer and genetic disorders. This includes extensive collaborations with biologists, geneticists and physicians.

TRANSPORTATION OPTIMIZATIONS

We recently initiated a sponsored research project with UPS. The project uses UPS data, along with machine learning techniques and algorithms to improve UPS delivery capabilities and develop a partnership to explore and develop cutting-edge techniques in the advanced analytics domain for the transportation and logistics industry.

IDENTIFYING RELATED DOCUMENTS

New Jersey Transportation Planning Authority has developed a system where transportation planners input documents, recommendations, and suggested improvements. We use Natural Language Parsing, Machine Learning and Topic Models to identify related documents. The approach provides a richer, sematic linkage as opposed to keyword-based searches.

USMAN ROSHAN
Associate Professor
Bioinformatics, machine learning, high performance computing.
usman.w.roshan@njit.edu

ZHI WEI
Associate Professor
Advanced data analytics: data mining, statistical modeling, machine learning.
zhi.wei@njit.edu

Senjuti Basu Roy

DATA SCIENCE WITH HUMAN IN THE LOOP

A traditional data science pipeline involves multiple steps — such as data cleaning, data exploration, data annotation, model building and result interpretation. Current best practice requires repetitive involvement of domain experts in each of these steps and is primarily trial-and-error based, a slow and costly process. Our objective is to optimize and automate data science pipelines by using crowdsourcing with citizen scientists and volunteers (high school students). We deal with multiple computational and algorithmic challenges, such as how to judiciously involve humans in the computational loop, provide assistance to researchers, adjust for human biases, errors and uncertainty in responses, as well as how to incorporate human feedback in the computational loop.

PROMOTING EFFECTIVE GROUP COLLABORATION

Drawing on research in social psychology, organizational theory and cyber-human systems, we address the challenges of next-generation collaborative crowdsourcing applications. This multidisciplinary research is developing unique mathematical models and algorithms to facilitate the creation of working groups, allowing them to make collaborative decisions while pursuing multiple objectives. The project also investigates how to enable feedback from humans during collaboration and combine that feedback with other data for dynamically determining evolving preferences.

SENJUTI BASU ROY
Assistant Professor
Database management, data analytics, algorithm courses.
senjuti.basuroy@njit.edu

Shaohua (David) Wang

MINING CONCISE AND USEFUL TIPS FOR PROGRAMMING FROM HETEROGENEOUS DATA SOURCES

A tremendous amount of data containing coding tips or suggestions exists on a wide range of data sources, such as YouTube, Q&A websites, reviews, etc. Mining such tips can help software practitioners, such as developers and students, grasp aspects of learning and using a programming language. We are developing deep learning techniques to accurately classify and extract small, useful and concise coding suggestions from crowdsourcing websites.

NEXT GENERATION QUESTION AND ANSWER SYSTEM FOR SOFTWARE DEVELOPERS

Typically, a software developer's question contains long text, code examples, "software semantics" and diagrams. This data diversity hinders many state-of-the-art techniques and systems perform very poorly when automatically answering such questions. We are developing AI techniques that analyze developers' questions and read documentations using deep learning techniques to compose accurate answers.

SHAOHUA WANG
Assistant Professor
Software engineering, artificial intelligence lab, machine learning.
shaohua.wang@njit.edu

DATA SCIENCE

IMAGE AND VIDEO

Brook Wu

EARLY DETECTION OF FAKE NEWS ON SOCIAL MEDIA

In the midst of today's pervasive influence of social media, automatically detecting fake news has become an important, but challenging task. We have developed a unique system that detects fake news based on user characteristics of news spreaders, e.g., people who tweet and retweet a news article. Approaches that rely on machine learning algorithms may not detect fake news early enough because the data required for detection is often insufficient at the early stage of news propagation. Experimental results show that our approach can effectively detect fake news very early in the process, e.g., 90 percent accuracy given 10 retweets and 95 percent accuracy given 60 retweets.

Our faculty are actively pursuing research in image and video analysis. Both NSF and U.S. DOT have funded projects in this area. Our research features improvements over the state-of-the-art in image analysis, as well as in applications to transportation problems.

BROOK WU
Associate Professor
Text mining, information extraction, information retrieval.
yi-fang.wu@njit.edu

ANALYSIS

Frank Shih

ELIMINATING THE BLIND SPOT

Car video recorders can capture scene information with different modalities, angles, resolutions and lens sensors. We are working on a novel approach to stitch together heterogeneous images, allowing a driver to view a scene without blind spots. Research also includes developing feature detection and matching tools that will integrate multiple images from multimodal sensors and stitch them into a wide-angle image.

A DEEP LEARNING FRAMEWORK FOR IMAGE MORPHOLOGY WITH APPLICATIONS TO ECOLOGY AND CONSERVATION

We are developing image watermarking techniques that hide secret "digital watermarks" for copyright protection, image authentication, data privacy and broadcast monitoring. Using deep learning, we developed the highest capacity and lowest image distortion by using saliency detection and shape decomposition schemes for automated watermark embedding. We have achieved a very high testing accuracy rate of 98 percent on the "Bee Wings" image dataset, a significant improvement over traditional 80 percent rates.

Chengjun Liu

MAKING INTELLIGENT TRANSPORTATION SYSTEMS SMARTER

The New Jersey Department of Transportation (NJDOT) has designated more than 400 CCTV video cameras distributed throughout the state to perform 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 advanced 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 incident detection and monitoring. We have investigated and developed fully-automated video analytics systems to replace human operators for automated traffic incident detection and monitor e-cameras installed along the major New Jersey highways. We test the proposed technologies and benchmark their performance through online testing using NJDOT test cameras.

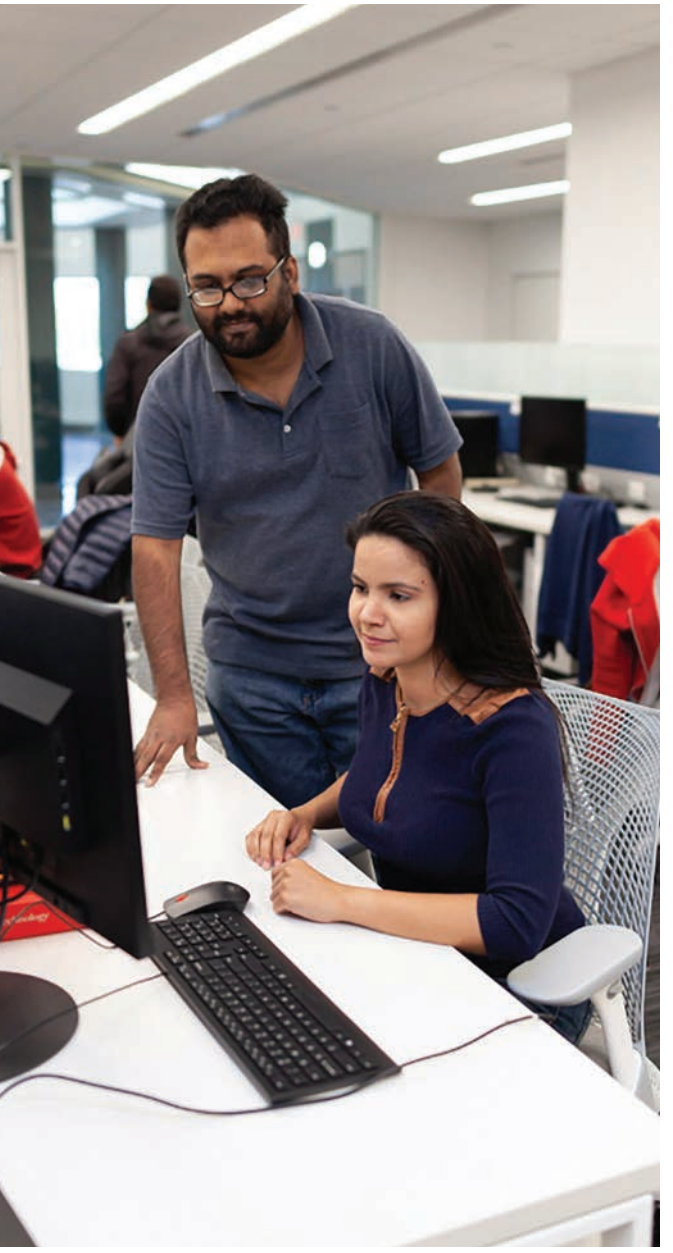
FRANK SHIH
Professor
Digital watermarking, digital forensics, robot sensing, neural networks.
frank.y.shih@njit.edu

CHENGJUN LIU
Associate Professor
Computer vision, pattern and face recognition, video processing.
chengjun.liu@njit.edu

DATA AND SOFTWARE SECURITY

The security research group has done extensive work in homomorphic encryption, supply chain security, blockchain and device/application security. It also makes significant contributions to open source software.

Kurt Rohloff received the first DARPA Young Faculty Award at NJIT. Initial results of a project in cliptography led to a conference award (top 3 rated) paper at Asiacrypt.



Kurt Rohloff

COMBATING DATA LEAKS: RAMPARTS, PARAPET, AND REVET

We work on a trio of projects that develop and apply software engineering tools to a new family of encryption technologies. These projects will result in a general open-source software library allowing organizations to outsource computation to cloud computing environments, without risking privacy and leaking sensitive information to potential adversaries.

IMPROVING USABILITY OF OPEN SOURCE SOFTWARE

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

KURT ROHLOFF
Associate Professor

Encrypted computing, lattice encryption implementation, homomorphic encryption, cryptographic program optimization.
rohloff@njit.edu

Qiang Tang

BLOCKCHAIN TECHNOLOGY AND ITS APPLICATIONS

Our research focuses on advancing three areas of blockchain technology, i.e., decentralized applications, better consensus protocols and underlying cryptographic building blocks. This includes building anonymous, yet accountable, applications on top of (imperfect) blockchain platforms and light client-friendly blockchain applications (such as for mobile phones and the internet of things). The results include abstraction and construction of new cryptographic primitives.

CLIPTOGRAPHY: CRYPTOGRAPHY AGAINST SUBVERSION ATTACKS

Software or hardware implementations (including open source code) of cryptographic functionalities may deviate from formal specifications. We are designing a new generation of cryptographic specifications that preserve security even in a malicious implementation or situation that contains backdoors. With this design, the subverted implementation can be detected via black-box testing.

QIANG TANG
Assistant Professor
Big data analytics; cyber security; parallel and distributed computing.
qiang@njit.edu

DATA AND SOFTWARE SECURITY

Iulian Neamtiu

IMPROVING SMARTPHONE RELIABILITY AND SECURITY

We have developed a wide range of approaches to improve smartphone reliability and security, including static and dynamic program analyses, record-and-replay systems, runtime systems for Moving Target Defense, app self-healing and automatic test generators. These approaches have been released as open source code and have found security and reliability issues in many popular apps, such as NPR News, AirBnB, Waze and Facebook.

PROTECTING ARMY NETWORKS

As part of the Cybersecurity Collaborative Research Alliance, a joint effort between the U.S. Army Research Laboratory, Applied Communication Services and six universities, our research aims to advance the foundations of cybersecurity in the context of military networks. We are using a two-pronged approach: theoretical foundations of, and practical approaches for agile defense of Army devices and networks, as well as measuring and predicting socio-cognitive factors involved in human decision-making during cybersecurity engagements.

IULIAN NEAMTIU
Associate Professor

Programming languages, software engineering, and the smartphone side of systems/security.
iulian.neamtiu@njit.edu

Reza Curtmola

DEFENDING SOFTWARE SUPPLY CHAINS AGAINST HACKERS

We are developing "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 have been 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.

REZA CURTMOLA
Associate Professor

Cybersecurity, applied cryptography, security of cloud services, software supply chain, wireless networks.
reza.curtmola@njit.edu

BIG KNOWLEDGE AND MEDICAL INFORMATICS

Faculty research has been applied to a number of important applications. Yehoshua Perl and James Geller are Fellows of the American College of Medical Informatics that have worked on improving medical ontologies. An ontology is one of the mechanisms used to encode diagnosis in electronic health records.

Applications to find parking in cities, as well as relieve traffic congestion are other examples of research results being applied to real world problems.

Results have been published in AMIA, the top U.S. conference in Medical Informatics and in JBI, one of the top journals.

Yehoshua Perl, James Geller and Michael Halper

SUMMARIZING, VISUALIZING AND CORRECTING LARGE COLLECTIONS OF MEDICAL TERMS

Biomedical professions have collected large repositories of medical terms covering diseases, diagnoses, drugs, anatomy, bacteria, genes, chemicals, medical procedures, and more. We have 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. 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.

YEHOSHUA PERL
Professor
Medical informatics, semantic web and ontologies.
yehoshua.perl@njit.edu

JAMES GELLER
Professor
Medical informatics, semantic web technology, object-oriented database modeling, knowledge representation.
james.geller@njit.edu

MICHAEL HALPER
Professor
Controlled terminologies, ontologies, object-oriented databases, conceptual modeling, open data models.
michael.halper@njit.edu

APPLICATIONS

Grace Wang

DECENTRALIZED CROWDSOURCING

Previous work on crowdsourcing has resulted in many clever incentive mechanisms, but all are enforced via a trusted central platform. We focus on establishing the theoretical foundation of decentralized crowdsourcing based on blockchain technology, as well as novel applications of these systems.

PREDICTING PASSENGER DEMAND USING DEEP "SPATIO-TEMPORAL FUZZY NEURAL NETWORK"

Predicting passenger demand in public transportation is a challenge, since this is determined by spatial and temporal factors, and other random factors, such as weather. We are developing a "Spatio-Temporal Fuzzy Neural Network" to accurately predict passenger demand in the near term. The network is composed of stacked Convolutional Long Short-Term Memory, fuzzy neural networks, convolutional layers and fully connected layers. Extensive experiments on real data show that this network outperforms state-of-the-art approaches.

EASING TRAFFIC CONGESTION AND COMMUTER CLOG

We are developing a cutting-edge, self-cognizant urban traffic signal control system based on machine learning technologies. This responsive signal control system adapts to changing traffic conditions.

Cristian Borcea

PREDICTIVE ONLINE ADVERTISEMENT SYSTEMS

Online advertising is a multi-billion dollar industry. We aim to improve online ad efficacy and reduce ad-annoyance of users through predictive machine learning algorithms. We propose probabilistic latent class models that predict the viewability (in the technical sense of online advertising) of any given scroll depth for a user-page pair and deep learning models to predict the viewability of any page depth for any given user dwell time. Currently, we are studying techniques to find a balance between publisher's revenue and user experience in the context of the ad-blocking battles between users and web publishers.

FINDING FREE PARKING IN CITIES

Finding an available free parking space in a city during peak hours is challenging. We have created a system for assigning free curbside parking spaces to drivers in cities that reduces driving time to the parking spot and walking time from the parking spot to the destination. Currently, we are designing a distributed system for free parking assignment that is scalable and protects the driver's privacy.

GRACE WANG
Professor
Mobile computing, deep learning applications, blockchain technologies, intelligent transportation.
guiling.wang@njit.edu

CRISTIAN BORCEA
Professor
Mobile computing and sensing, ad hoc and vehicular networks, cloud and distributed systems.
cristian.m.borcea@njit.edu

CLOUD, HIGH PERFORMANCE COMPUTING

Researchers are investigating multiple aspects of cloud computing. Reducing latency is one of the key aspects of the research and it includes use of parallelism and algorithms for scheduling.

A key result of this research is enabling applications to fully leverage the computing capability of hardware and devices to achieve optimal data processing speeds. Research results, such as APPLES – a solution for reducing spin-lock overhead – have attracted interest from leading companies such as Tencent and Didi.

Alex Gerbessiotis

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 are necessary to increase processing speed and improve memory performance.

PARALLEL COMPUTING TECHNIQUES IN SEQUENTIAL (AKA 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.

ALEX GERBESSIOTIS
Associate Professor

Architecture-independent parallel algorithm design and implementation.
alexg@njit.edu

CLOUD, HIGH PERFORMANCE COMPUTING

Xiaoning Ding

VIRTUALIZATION OF HETEROGENEOUS AND NON-UNIFORM MEMORY HIERARCHIES

Accessing data saved in heterogeneous and non-uniform memory hierarchies is an increasingly important factor for the performance of many applications. However, in the cloud, system software, particularly virtualization software, causes data-intensive applications to suffer a significant performance penalty. We improve memory virtualization technology to build virtual memory hierarchies. Virtual memory hierarchies have similar architectures and features as those of new memory systems. Thus, they can effectively serve as the portal for applications in virtual machines to efficiently access the data in new memory systems.

SYSTEM SOFTWARE FOR SCALABLE COMPUTATION IN THE CLOUD

As the resources available for computation keep increasing on today's computers (e.g., multi-core/many-core processors and accelerators), 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.

XIAONING DING
Associate Professor
Cloud computing infrastructures, parallel and distributed systems, operating systems, database systems, computer architecture.
xiaoning.ding@njit.edu

Jing Li

SCHEDULING FOR INTERACTIVE CLOUD SERVICES

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

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 from 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 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.

Andrew Sohn

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 problems and improve cloud servers, such as Amazon's Elastic Compute Cloud and Microsoft Azure, and meet the power and computing requirements of mobile, as well as 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 large-scale 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 graphs that change over time as one has to balance partition quality and execution time. We have 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.

JING LI
Assistant Professor
Real-time systems, parallel computing, cyber-physical systems, interactive online cloud services.
jing.li.2@njit.edu

ANDREW SOHN
Associate Professor
Adaptive computing infrastructure, autonomous migration of Linux virtual machines, maximizing cluster utilization.
andrew.sohn@njit.edu

AUGMENTED REALITY, VIRTUAL REALITY AND GAMING

Research includes cooperative computing between humans and computers, as well as using gaming and virtual reality to enhance education. Projects also include using gaming for data collection.

Eric Nersesian

CREATING IMMERSIVE CLASSROOM ENVIRONMENTS

We explore the integration of immersive technologies in classroom environments to increase student comprehension and engagement. Modern educational approaches rely on a combination of multitasking and abstract reasoning. Interactive multimedia educational technologies have been documented to have marked improvements over traditional educational media.

Educational technologies offer significant benefits to learners, enhancing visual short-term memory, spatial cognition, multitasking and executive function. Additionally, there is an increasing need to develop effective multimedia content for student engagement. As each generation becomes more accustomed to rich multimedia experiences in everyday facets of their lives, the delivery mechanisms of their educational content need to keep pace and supplement the traditional methods of lecture and textbook-based instruction with technological alternatives.

Frank Biocca

DESIGN OF VIRTUAL ENVIRONMENTS AND BRAIN SENSING FOR AUGMENTING HUMAN COGNITION

Virtual reality and augmented reality hardware and applications are deployed with the goal of significantly affecting user engagement, mental performance, or social processes. We explore how the design features of virtual and augmented reality environments influence or augment human cognition. Research examines the design of virtual and augmented reality hardware, software interaction techniques, and applications to determine the effects on different aspects of cognitive performance such as human perception, sense of presence, improved learning, higher engagement, spatial cognition and social cognition. Untethered brain sensors and psychophysiological measures (i.e., functional Near Infrared Spectroscopy) provide real-time measures of user cognition in a virtual or augmented reality system. Brain sensing (i.e., neurocognitive indicators) are also used as user input for brain adaptive virtual environments.

ERIC NERSESIAN
Lecturer

Interactive computer graphics for training and educational applications.
eric.nersesian@njit.edu

FRANK BIOCCA
Professor

Virtual and augmented reality systems, components for brain-computer interfaces, real-time public opinion measurement.
frank.a.biocca@njit.edu

Michael Lee

ADVANCING STEM EDUCATION WITH GAMING

We are exploring unique approaches to teaching STEM (Science, Technology, Engineering and Math) topics in formal and informal learning environments. Our work on "Gidget" (helpgidget.org), an online game to teach students introductory programming concepts, has reached thousands of people all over the world. Gidget has shown to be effective in attracting and engaging a broad audience, including women and underrepresented minority groups in computing. In addition, we are working to increase access to computing, and we are participating in efforts to provide access to computing education, resources and mentors for local K-12 youth. We are also working with Oculus Education to support STEM education in a local high school using virtual reality (VR) headsets.

Amy Hoover

HUMAN COMPUTER COLLABORATION

We explore how humans and computers can work together to solve problems by conceptualizing creativity as a search in a structured space of computational artifacts. Our work has been applied to understanding game design, assessing learning in educational games, creating game content and behavior, as well as facilitating human creativity.

MICHAEL LEE
Assistant Professor
Human-computer interaction (HCI) and computing education research (CER).
michael.j.lee@njit.edu

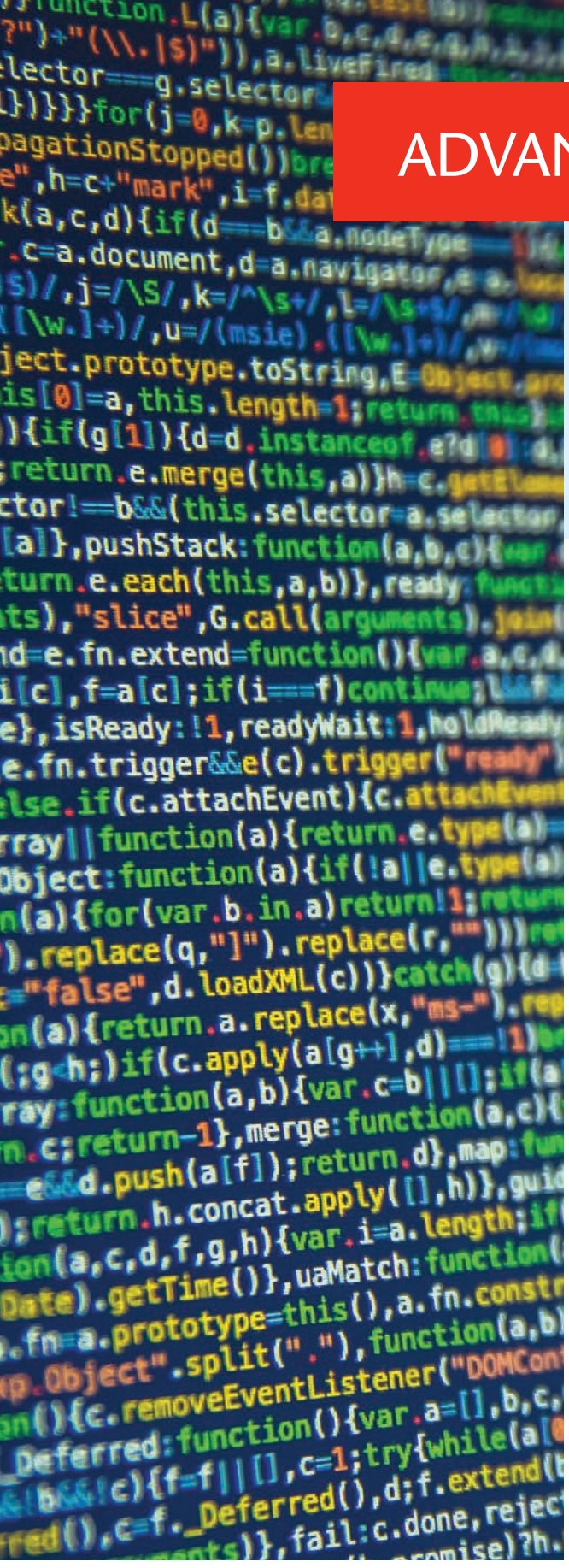
AMY HOOVER
Assistant Professor
Data science, AI-based human-computer interaction, artificial intelligence.
amy.hoover@njit.edu

Cristian Borcea

USING GAMIFICATION TO COLLECT DATA

Measuring and harvesting large datasets from the field using sensor devices, known as mobile crowdsensing, can identify a broad spectrum of environmental, infrastructure and social needs. However, developing a method to incentivize people to collect and share sensor data in some settings remains a major challenge. We have designed, built and evaluated two Android-compatible crowdsensing systems for NJIT students: a mobile game that uses an incentive-centered design to convince participants to cover all the regions of a target area, and a micropayment structure that allows users to pick their own sensing tasks. The results suggest that gaming is a cost-effective solution for uniform area coverage, while micropayments work well for sensing tasks with tight time constraints or long-term tasks for personal analytics.

CRISTIAN BORCEA
Professor
Mobile computing and sensing, ad hoc and vehicular networks, cloud and distributed systems.
cristian.m.borcea@njit.edu



ADVANCED ALGORITHMS

There is a wide variety of formal methods and algorithmic research at NJIT. The work includes optimization, risk assessment, numerical algorithms and program repair.

Researchers publish in top-tier conferences and have received recognition, including a Best Paper award at ICALP 2017 and an NSF CAREER award.

Ioannis Koutis

NUMERICAL ALGORITHMS

We aim to design very fast and robust numerical algorithms for fundamental computational problems, including linear system solvers. These algorithms use mathematical insights from the spectral analysis of graphs, which can be applied to attack data mining and machine learning problems on large networks. We also design parameterized algorithms for difficult and complex optimization problems. We pioneered the method of "algebraic fingerprints," a term we coined for a general method that encodes combinatorial problems as polynomials, with each possible solution having a monomial as its "algebraic fingerprint." This research has led to breakthrough results for classical algorithmic problems, such as the Hamiltonian Cycle problem and the single exponential time algorithms for problems parameterized by tree width.

IOANNIS KOUTIS
Associate Professor
Fast linear system solvers, spectral graph algorithms, machine learning, data mining.
ioannis.koutis@njit.edu

Ali Mili

CONVERGENCE: INTEGRATING TERMINATION AND ABORT FREEDOM

The condition under which a program terminates and the conditions under which a program executes without causing an abort (division by zero, array reference out of bounds, nil pointer reference, etc.) are usually analyzed separately. We argue that they should be analyzed jointly to produce more complete results. We develop theoretical foundations for analyzing these two aspects jointly and use them to develop an automated tool to this effect.

REDUNDANCY: THE MUTANTS' ELIXIR OF IMMORTALITY

Determining whether a mutant (a program obtained by applying an elementary syntactic modification to a base program) is equivalent to a base program is known to be undecidable; in practice, determining this is tedious, error prone and non-scalable. We address this problem by estimating the number of equivalent mutants using redundancy metrics. For most applications of mutation testing, it is not necessary to identify equivalent mutants; it is sufficient to estimate their number. We have shown a significant statistical correlation between the redundancy of a base program and the ratio of equivalent mutants that it is prone to generate, and we are developing a tool that estimates the number of equivalent mutants of a base program using these statistical relations.

A GENERIC ALGORITHM FOR PROGRAM REPAIR

Program repair usually involves two phases: patch generation, when repair candidates are generated from the original by some tentative transformations; and patch validation, when repair candidates are evaluated to select a valid repair. Whereas patch validation usually proceeds by optimizing some fitness function, we argue that it is best carried out using relative correctness. The repair must be more correct than the original program. We are implementing a generic repair algorithm on the basis of this premise and evaluating its performance on standard benchmarks.

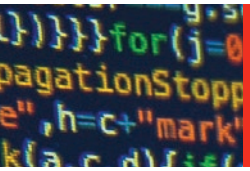
Marvin Nakayama

EFFICIENT COMPUTATIONAL METHODS FOR RISK ASSESSMENT

We aim to achieve a substantial reduction in the number of errors in Monte Carlo simulation, which is a computational technique used to measure risk in quantitative analysis and decision-making. Having the ability to precisely estimate the probability of large losses in financial portfolios or assess the likelihood of a hurricane severely damaging critical infrastructure benefits professionals in widely disparate fields as finance, civil and mechanical engineering, and project management.

ALI MILI
Professor
Software engineering, high assurance systems engineering, discrete mathematics, software metrics.
ali.mili@njit.edu

MARVIN NAKAYAMA
Professor
Simulation modeling and analysis, risk analysis, energy modeling, fault-tolerant systems.
marvin@njit.edu



ADVANCED ALGORITHMS

Baruch Schieber

DYNAMIC GRAPH ALGORITHMS

Graphs are used extensively to model various kinds of networks, like transportation networks or social networks. In most of real life applications these networks change over time and thus their characteristics are changing as well. The goal of dynamic graph algorithms is to compute these characteristics over time as efficiently as possible. Clearly, 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. MIS algorithms constitute a useful subroutine for breaking. The MIS problem has extensive connections to a plethora of fundamental combinatorial optimization problems such as maximum matching, minimum vertex cover and graph coloring. We have developed several dynamic algorithms for computing MIS, among them 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 several applications, e.g., 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, and thus 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 indeed 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.

James Calvin

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.

BARUCH SCHIEBER
Professor
Algorithms, mathematics of artificial intelligence (AI), and optimization.
baruch.m.schieber@njit.edu

JAMES CALVIN
Professor
Global optimization, probabilistic analysis of algorithms.
james.m.calvin@njit.edu

SOCIAL/ MOBILE/ EDUCATION

This research focuses on how technology can help in social environments and the classroom. For instance, a grant from the Mozilla Foundation examines how female and LGBTQ live streamers cope with online harassment.

Michael Bieber

REDEFINING STUDENT ENGAGEMENT

Participatory learning deepens knowledge through active involvement in the entire lifecycle of assignments, exams and other course activities — including peer grading and self-assessment. We are developing a web application that functions as an educational dashboard, integrating commenting and mentoring capabilities, and automatically assigning tasks to students.

WEARABLE TECHNOLOGY IN CLASSROOMS

Classroom dynamics continue to evolve as more students will use their personal wearable technology — watches, wristbands and contact lenses — in ways that are undetectable by instructors and peers, which would create an atmosphere of distrust. We are the first to systematically investigate and determine how best to design learning, teaching and assessment when personal wearable technology is used undetectably in classroom settings.

Brook Wu

AUTOMATING MEANINGFUL LEARNING

We are developing a system that will improve students' understanding of course content. As a pedagogical strategy, "Write-and-Learn" provides a framework that generates automated formative feedback by comparing the concepts introduced in teaching materials and students' writing assignments. Our work will evaluate the effectiveness of automated formative feedback and explore how such feedback can scaffold and promote meaningful learning.

MICHAEL BIEBER
Professor
Web engineering, link-based services, hypermedia, community informatics.
bieber@njit.edu

BROOK WU
Associate Professor
Text mining, information extraction, information retrieval.
yi-fang.wu@njit.edu

SOCIAL/ MOBILE/ EDUCATION

Yvette Wohn

EXPLORING THE LINK BETWEEN TECHNOLOGY AND WELL-BEING

We are exploring the connection between health and technology usage — with an emphasis on the role social media play in facilitating social support and psychological well-being. We have developed mobile applications for first-year college students and women in STEM fields and are working on a system for those coping with opioid dependency. Other projects include looking at the role of temperature on feelings of closeness with conversational agents and the usage of wearable devices for health monitoring.

DESIGNING SYSTEMS FOR STROKE REHABILITATION

Games can be used to make physical therapy more enjoyable for stroke patients. With medical partners at Rutgers University, we are designing a system that will help both clinicians and patients utilize games for physical therapy.

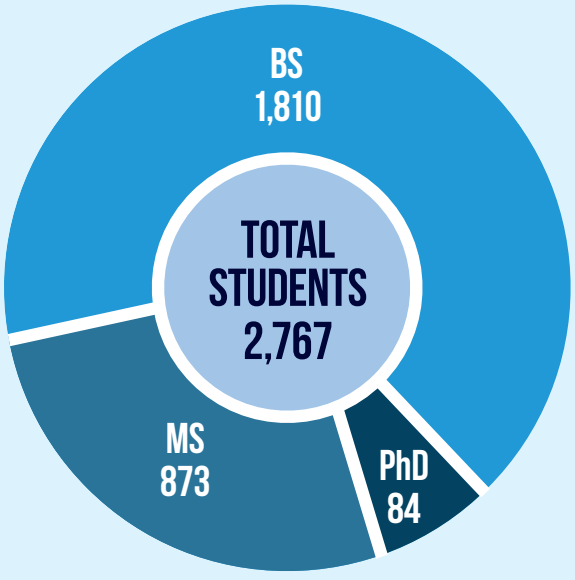
FIGHTING NEGATIVITY ONLINE

With burgeoning content online, how do we handle negativity, which ranges from cyberbullying to the posting of deviant content? We examine the influence of systems, algorithms and people on content moderation in collaboration with researchers from Google and Facebook. One of the projects examines how female and LGBTQ live streamers cope with online harassment.

YVETTE WOHN
Assistant Professor

Content moderation practices, teamwork in e-Sports, social media usage and wellbeing.
donghee.y.wohn@njit.edu

COLLEGE OF COMPUTING FACTS AND FIGURES



DEGREE PROGRAMS

Undergraduate

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

Graduate

M.S. Computer Science
M.S. Information Systems
M.S. Bioinformatics
M.S. Computing and Business
M.S. Business and Information Systems
M.S. Software Engineering
M.S. IT Administration and Security
M.S. Cyber Security and Privacy
M.S. Data Science

Doctoral Degrees

Ph.D. in Computer Science
Ph.D. in Information Systems

FACULTY

Tenure-Track
42

Instructional Lecturers
27

DIVERSITY

20%
of Students are Women

21%
of Students are Underrepresented Minorities

ALUMNI

Number of Annual Graduates
800

Average Starting Salary
$73,748

Ying Wu College of Computing
New Jersey Institute of Technology
University Heights
Newark, NJ 07102
computing.njit.edu

1M 1/19