RESEARCH REPORT
YING WU COLLEGE OF COMPUTING SCIENCES
SPRING 2016

CONTENTS

Research Report of the Ying Wu College of Computing Sciences Spring 2016

INTRODUCTION

"The pages of our math books should be covered with blood," said Renaissance man Alan Kay, one of the masterminds at Xerox PARC, at the IJCAI conference in Los Angeles in 1985, where he gave the keynote address on the challenges of being a scientist.

Kay's pronouncement was the conclusion of his musings about the struggles of doing scientific research. He noted that the clean results on the pages of a book do not reveal the blood, sweat, and tears—nor even hint at the many false starts, cul-de-sacs, and dead ends—suffered by scientists during long nights spent searching for answers.

It was the second scientific conference that I attended, the first where I presented a paper, and I was electrified. So impressed was I by Kay's address that I bought a cassette tape of his lecture (Remember what a tape is?) and memorized parts of it.

What Kay did not mention were the human challenges the young researcher is up against. Getting a paper published, even if the results are phenomenal, often requires several revisions, sometimes because of a reviewer who fails to appreciate the significance of the work. Once a scientist has decided what to say, it takes persistence, grit, and a certain mental flexibility to get the message out.

If publishing in a disinterested world is hard, getting financial support is even harder. While the number of Ph.D. level researchers is constantly growing, funding is, at best, remaining flat. Science and research in the twentieth century made the United States the most dominant economic and military power in history. Yet, there are not enough funding resources to continue this trend, and faculty need to compete for those that are available.

In the following pages you will read about the imaginative work of a group of intrepid researchers who have struggled to build beautiful edifices in abstract spaces, such as the Cloud. Their efforts have garnered financial support from the federal government, from companies, and from nonprofits. They are reaping the rewards of persistence through countless sleepless nights and struggles.

They are members of the Ying Wu College of Computing Sciences at New Jersey Institute of Technology, and I could not be more proud of their achievements.

Let me take this opportunity to thank Dean Marek Rusinkiewicz for many years of friendship and for three years of superb leadership of the college. We gratefully acknowledge Mr. Ying Wu for his interest and investment in the newly named Ying Wu College of Computing Sciences.

James Geller
Professor, Computer Science Department
Associate Dean for Research, Ying Wu College of Computing Sciences
New Jersey Institute of Technology

BIG DATA

Robust Network Fusion Algorithms for Detection and Localization of Radiation Sources In support of national defense, Domestic Nuclear Detection Office's (DNDO) Intelligent Radiation sensor Systems (IRSS) program developed networks of radiation counters for detecting, localizing and identifying low-level, hazardous radiation. This project is focused on (i) development and maintenance of key measurements, datasets, and benchmarks; and (ii) design, analysis and testing using DNDO IRSS datasets, of next-generation network fusion algorithms for detection and localization with robust performance across a wide range of scenarios. We are distilling DNDO IRSS datasets to generate key measurements and benchmarks, which will be released in two forms, one for public research and development and other for government use only, with approval from DNDO. Our detection approach, which is novel and somewhat counter-intuitive, contrasts sharply with current "detect first and then localize" methods. In our method, strategically selected small subnets are used to generate multiple source estimates, and their convergent clusters are fused to generate a robust source estimate. This source estimate is then plugged into sharpened probability ratio tests with adaptive thresholds to infer detection. In addition to being solidly founded in our theory and analytical results, this approach is supported by critical inferences from DNDO IRSS datasets. In several tests, the source estimates rapidly converged and stabilized at time-scales much faster than detection times.
NJIT Faculty: Chase Wu Funding: Department of Homeland Security (DHS)/Domestic Nuclear Detection Office (DNDO)

Technologies and Tools for Synthesis of Source-to-Sink High-Performance Flows A number of Department of Energy (DOE) science applications, involving exascale computing systems and large experimental facilities, are expected to generate large volumes of data, in the range of petabytes to exabytes, which will be transported over wide-area networks for the purpose of storage, visualization, and analysis. To support such capabilities, significant progress has been made in various components, including the deployment of 100 Gbps networks with future 1 Tbps bandwidth, increases in end-host capabilities with multiple cores and buses, capacity improvements in large disk arrays, and deployment of parallel file systems such as Lustre and GPFS. High-performance source-to-sink data flows must be composed of these component systems, which require significant optimizations of the storage-to-host data and execution paths to match the edge and long-haul network connections. The main objectives of this project are twofold: (1) develop and test component technologies and their synthesis methods to achieve source-to-sink high-performance of 100 Gbps traffic flows that can scale to 1 Tbps; and (2) develop tools with simple interfaces to users and applications that automate these capabilities by profiling the hosts summarized. The major research activities include the following: (a) synthesis methods that compose complex source-to-sink flows using individual storage and network flows, realized across the storage and local networks as well as long-haul connections; (b) automated tools that combine the synthesis modules with profiling modules for hosts, disks, and file systems, and network connections to provide the transport capabilities transparently to science users and applications.
NJIT Faculty: Chase Wu Funding: Department of Energy (DOE) Office of Science, Big Data-Aware Terabits Networking

An Integrated Approach to Performance Modeling and Optimization of Big-Data Scientific Workflows
Next-generation e-science is producing colossal amounts of data, now frequently termed Big Data, on the order of terabytes at present and petabytes or even exabytes in the predictable future. These scientific applications typically feature data- and network-intensive workflows comprised of computing modules with intricate inter-module dependencies. Application users often need to manually configure their computing workflows in distributed environments in an ad-hoc manner, which significantly limits the productivity of scientists and constrains the use of resources. We propose to build a three-layer workflow optimization architecture that seamlessly integrates three interrelated components: resource abstraction, module mapping, and node scheduling, based on rigorous algorithmic

design, theoretical dynamics analysis, and real network implementation, deployment, and evaluation. The successful completion of this project will (i) lay down a solid mathematical foundation for the analysis and control of system dynamics of distributed big data scientific workflows; (ii) produce a suite of cooperative mapping and scheduling optimization solutions to facilitate scientific collaborations; and (iii) add an additional level of intelligence to existing workflow engines widely adopted in the current grid and cloud computing middleware. NJIT Faculty: Chase Wu
Funding: National Science Foundation (NSF): Computer and Network Systems (CNS)

Big-Data Visualization The over operator is commonly used for blending in various visualization techniques. In the current form, it is a binary operator and must strictly follow a specific composition order of all participating operands, hence posing a significant performance limit. We derive a set of generic formulas for the over operator that work with any number of operands and completely remove the restriction on the composition order. We prove the correctness of these formulas and provide a step-by-step illustration in a blending context. We implement a sequential and a parallel version of the improved over operator and apply them to the image composition problem where operands are received out of order with different arrival time intervals. The performance superiority of the improved over operator over the original one is established by rigorous theoretical analyses and further validated by extensive experimental results. NJIT Faculty: Chase Wu

Green Networking The transfer of Big Data in various applications across high-performance networks (HPNs) in a national or international scope consumes a significant amount of energy on a daily basis. However, most existing bandwidth scheduling algorithms concern only traditional objectives such as data-transfer time-minimization, and very limited efforts have been devoted to energy efficiency in HPNs. We consider two widely adopted power models—power-down and speed scaling—and formulate two advance instant bandwidth scheduling problems to minimize energy consumption with data transfer deadlines under packet loss constraint. We prove the NP-completeness of both problems, and design an epsilon-approximation algorithm for the constraints using the power-down model and a polynomial-time optimal solution for a special case. We also design an approximation

algorithm and a heuristic approach that considers the tradeoff between objective optimality and time cost in practice for the problem using the speed-scaling model. The performance superiority of the proposed solutions in terms of energy saving is illustrated by extensive results based on both simulated and real-life networks in comparison with existing methods. NJIT Faculty: Chase Wu

Multi-Instrument Database of Solar Flares Solar flares represent the most prominent manifestation of the Sun's magnetic activity. They are accompanied by high-speed coronal mass ejections and high-energy particles, which play critical roles in the Earth's space environment and space weather. Solar flares are observed by several NASA space missions: IRIS, SDO, STEREO, Hinode, RHESSI, FERMI, SOHO, ACE, and WIND, as well as by many ground-based observatories, including BBSO and EOVSA observatories operated by NJIT. The flare emissions cover the whole range of the electromagnetic spectrum, from radio waves to gamma rays. Also, substantial amounts of energy are released in the form of accelerated particles and high-speed plasma flows. Currently, the flare data are collected in databases of individual instrument teams, and thus it is extremely difficult to automatically categorize them according to a user-defined set of common physical properties to collect these data and prepare them for multi-instrument, multi-wavelength analysis. The proposed work includes development of compact representations and effective, efficient, and scalable search algorithms. In addition, we are investigating machine-learning methodologies for automatic mining and classification based on both alpha-numerical data and visual features. NJIT Faculty: Gelu Nita, Alexander Kosovichev, Vincent Oria Funding: National Aeronautics and Space Administration (NASA)

CYBERSECURITY

CAREER: Secure and Reliable Outsourced Storage Systems Using Remote Data Checking This project establishes a practical remote data checking (RDC) framework as a mechanism to provide long-term integrity and reliability for remotely stored data. At the same time, the project seeks to develop new functionality for remote data checking that overcomes limitations of early RDC protocols and improves the usability and deployability of RDC on existing cloud storage infrastructures. Unlike previous work, this research takes a holistic approach and considers RDC protocols that minimize the combined security costs of all data management phases over the lifetime of a distributed storage system. This includes prevention, repair, and retrieval. Maintaining the health of the data in a distributed storage system requires various transformations to be applied to the data and requires data to migrate among storage servers. This project develops novel RDC protocols that are compatible with the full range of replication, erasure coding, and network coding operations employed by distributed storage systems, thus enabling owners to maintain better control over their data. This project increases the transparency of cloud storage platforms and improves the security dimension of storage outsourcing, enabling wider adoption of cloud storage technologies.
NJIT Faculty: Reza Curtmola Funding: National Science Foundation (NSF) Trustworthy Computing

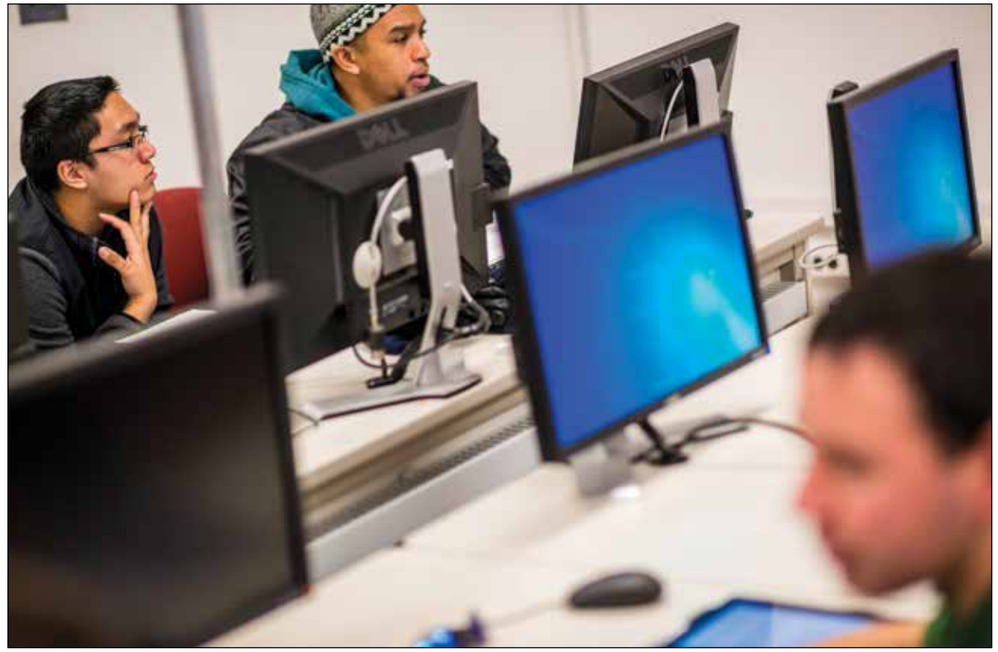
Toto: Securing Software Supply Chain Logistics We seek to investigate, design, and ensure adoption of a system that provides end-to-end guarantees about the software supply chain. We call this system "Toto," which comes from the Latin in toto meaning "on the whole." Toto protects software from the moment it is committed by a developer and ensures that the chain of trust can be followed all the way until users install the software. Toto generates metadata about each process in the software chain (e.g., cryptographically signed reports), and links together and carries these metadata throughout the entire chain. In particular, Toto provides a mechanism which allows end users to verify

that (1) the source code was not tampered with in the version control system; (2) the code was tested for bugs using code review, unit testing, and static and dynamic analysis tools; (3) the build system produced correct outputs; (4) the distribution and update infrastructure delivered authentic versions of the software; and (5) each step of the supply chain used validated inputs from the prior step. Since processes such as code review and testing practices are now publicly visible, Toto provides incentives for developers to follow safe software practices. NJIT Faculty: Reza Curtmola Funding: Defense Advanced Research Projects Agency (DARPA) Information Innovation Office

PALISADE This is a next-generation lattice encryption library designed to support fully homomorphic encryption, encrypted program obfuscation, and proxy re-encryption. The library is currently in use by multiple industry projects. The library is in active development with multiple federal funding sources in collaboration with multiple academic, industry, and government partners and co-developers. NJIT Faculty: Kurt Rohloff Funding: Defense Advanced Research Projects Agency (DARPA) Information Innovation Office

OPERA This is a multi-year, multi-million-dollar DARPA-funded research and development effort to demonstrate practical encrypted program obfuscation. NJIT Faculty: Kurt Rohloff Funding: Defense Advanced Research Projects Agency (DARPA) Information Innovation Office

PARAPET This is a federally funded research and development project to apply FHE technologies to protect against data exfiltration from commodity cloud-computing environments. This work is one of the first applications of



practical fatal hardware error (FHE) technologies and builds on the advancements from multiple prior DARPA-funded research and development projects. NJIT Faculty: Kurt Rohloff Funding: National Security Agency (NSA) National Centers of Academic Excellence in Information Assurance Education (CAE)

Cyber-Security Collaborative Research Alliance This is a joint effort among the Army Research Laboratory and five universities, whose goal is to advance the theoretical foundations of cyber science in the context of Army networks. The goal of this program is to understand and model the risks, human behaviors and motivations, and attacks against Army systems. Such understanding and models will lead to an asymmetric advantage in cyber domains against known and unknown attackers in the ability to detect and thwart attacks as well as allow mission progress in the face of ongoing and evolving threats. NJIT Faculty: Iulian Neamtiu Funding: Army Research Lab

Differential Types and Declarative Hypothesis Testing for Software Evolution Changes made to source code in the course of evolution might not have the intended meaning (fix bugs, add features), which negatively affects both software producers and software consumers, as well as having high economic costs. This work uses mathematical models of program behavior and empirical software engineering to better understand and facilitate the software evolution process. Our research has helped uncover certain kinds of errors before software is deployed and reduce the incidence of, and costs associated with, software bugs and incorrect software updates. NJIT Faculty: Iulian Neamtiu Funding: National Science Foundation (NSF)

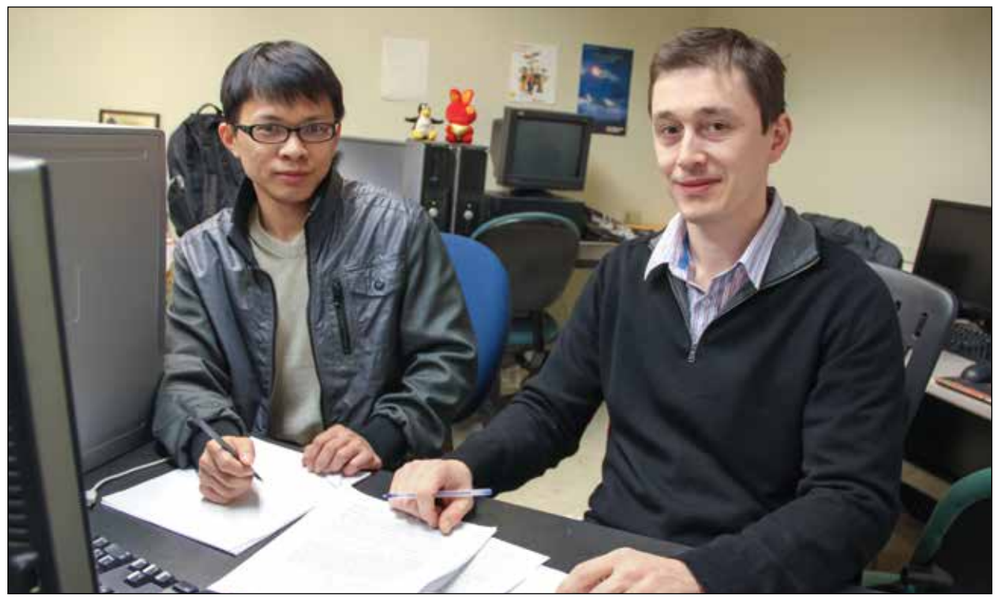
Program Analysis for Smartphone Application Security This project has developed a variety of analyses aimed at improving Android's reliability and security. First, we have showed that data-mining software repositories can reveal common classes of errors in mobile applications. Second, we have developed VALERA, a record-and-replay approach that helps developers or users with a variety of tasks, such as reproducing executions, finding and fixing concurrency bugs, and application profiling. Third, we have constructed analyses that have (a) found a new class of Android application errors we named "resume/restart errors" and (b) found and reduced the security risks posed by Android applications. NJIT Faculty: Iulian Neamtiu Funding: National Science Foundation (NSF)

NJIT Secure Computing Initiative (SCI) The NJIT Secure Computing Initiative (SCI) capitalizes on the Ying Wu College of Computing Sciences's broad range of strengths in cyber-security to train and place successive cohorts of expert cyber-security professionals in the federal, state, and local agencies. Specifically, the students funded by this grant receive three-year Scholarship for Service (SFS) funding that covers their last year in one of the B.S. programs offered by our college and two years in one of our two M.S. programs on cyber-security. Specifically these are an M.S. in cyber-security and privacy, and an M.S. in information technology administration and security. In addition to taking traditional courses, these students work in teams led by one of the principal investigators on out-of-class research projects drawn from areas such as cryptographic solutions to address various facets of secure data outsourcing to the Cloud, secure and efficient content distribution in wireless networks, trusted mobile people-centric sensing, privacy protection in mobile computing, and cyber-security quantification and decision-making. NJIT Faculty: Vincent Oria, Cristian Borcea, Reza Curtmola, Ali Mili Funding: National Science Foundation (NSF)

NETWORKING, CLOUD COMPUTING, AND MOBILE COMPUTING

Avatar: Mobile Distributed Computing in the Cloud Smart phones and tablets have become the personal devices of choice for most people. However, mobile-distributed applications are currently difficult to deploy because mobile devices are not always reachable and have limited resources, such as energy and bandwidth. The goal of this research is to make the promise of mobile-distributed computing a reality for our society to enable new beneficial applications in areas such as healthcare, safety, and social interaction. We have designed and implemented a system architecture in which a mobile user has one or more mobile devices and an "avatar" hosted in the Cloud. An avatar is a per-user software entity instantiated as a virtual machine in the Cloud, which acts as a surrogate for the user mobile devices, thus reducing the workload, storage resources, and bandwidth needed on the mobiles. Avatars make mobile devices smarter because they have a larger storage capacity, computing power, bandwidth, and availability than mobiles. Implicitly, they save energy on the mobiles and improve the response time for many apps. The avatars are available 24/7, even when their mobile devices are offline because of poor connectivity or because they are turned off. NJIT Faculty: Cristian Borcea, Narain Gehani, Xiaoning Ding, Reza Curtmola Funding: National Science Foundation (NSF) Computer Systems Research (CSR)

Crowdsensing in the Wild with Aliens and Micro-Payments Mobile crowdsensing can be used to enable a broad spectrum of applications, ranging from monitoring pollution or traffic in cities to epidemic disease monitoring or reporting from disaster situations. We recently designed, built, and evaluated through user studies two crowdsensing systems based on different types of incentives: a mobile game that uses in-game incentives to convince participants to cover all the regions of a target area, including the unpopular ones; and a micro-payment that allows users to pick the sensing tasks they want to execute according to their own criteria and provides a small payment for each task. Any



student on the NJIT campus was allowed to participate in these studies, and all they had to do in order to participate was to install our mobile apps on their Android phones. More than 50 students participated in each study, and the duration of the studies were 35 days for mobile gaming and two months for micro-payments. The analysis of the results suggests that gaming is a cost-effective solution for uniform area coverage, while micro-payments work well for sensing tasks with tight time constraints or long-term tasks for personal analytics. NJIT Faculty: Cristian Borcea, Reza Curtmola

Efficient and Scalable Cloud Infrastructures on New Computer Architectures Computer architectures change rapidly to satisfy the increasing demand for computing power from data-intensive applications, such as Big Data computing. Cloud-computing platforms form the backbone of the modern computing infrastructures. New changes in computer architecture require that the system software for Cloud computing be continually improved, so that computation in the Cloud can benefit from these changes to achieve high performance, high efficiency, and low energy consumption. The project targets the foundation of Cloud computing, virtualization, and builds new virtualization technology to manage the new computer architectures and hide the complexity for the computation in the Cloud to achieve efficient utilization. Currently the project focuses on (1) building highly scalable virtual machines and containers to leverage accelerators and the increasing core count in CPUs, and (2) the virtualization of new memory devices and new memory technology, such as NVM. NJIT Faculty: Xiaoning Ding

Smart Bus System Under Connected Vehicles Environment This research advances a smart bus system (SBS) powered by bus-to-devices wireless communications technology, including, but not limited to, 3G, 4G/LTE, Wi-Fi, dedicated short-range communications (DSRC), and Bluetooth. SBS is an innovative urban bus operation system integrated with information technology (IT) to enhance the efficiency of bus operation, to encourage bus ridership, and to improve the mobility and sustainability of urban transportation. SBS enables (1) a bus to take a shorter path to catch up when it is behind the schedule and no passengers need to get off or get on and (2) passengers to send a hold request so that the transfer bus, if at all possible, can wait for the transfer passenger. NJIT Faculty: Grace Wang Funding Agency: University Transportation Research Center (UTRC)

Autonomous Migration of Containers and Virtual Machines At the bottom of Cloud-computing infrastructures is an army of virtual machines and containers running on LAMP-based physical machines. A challenge in providing on-demand Cloud services is the use of resources and energy. Imbalance of loads among the physical machines leads to inefficiency, which can result in slow-to-no response of the Cloud. A solution to load imbalance is live migration of virtual machines and containers. This research focuses primarily on kernel-level migration of virtual machines and containers autonomously without human intervention in such a way that when migrated, the state of the Cloud is back to normal. NJIT Faculty: Andrew Sohn

BIOMEDICAL INFORMATICS

ExploreNet The goal of ExploreNet is to develop a suite of algorithms, tools, and web servers for inferring biological and social networks using graph-mining algorithms. In a collaboration among colleagues, graduate and undergraduate students, we adopt a five-pronged approach to the development of the ExploreNet framework. Specifically we design, develop, and implement new software for (1) reconstructing networks using a pattern

discovery approach; (2) inferring networks using weakly-supervised and semi-supervised learning methods; (3) predicting missing links in integrated, heterogeneous networks; (4) identifying structural differences in networks by extending our previous algorithms for RNA structure alignment and tree pattern matching; and (5) inferring networks on Clouds. Taken together, we expect that the ExploreNet framework will significantly contribute to advancing the development of computational approaches for network inference and analysis. The techniques resulting from this project will assist in data categorization in social networks. They will also be useful for biological network annotation, which helps scientists gain a better understanding of the functional elements and regulatory circuits of cells. The software tools developed from this project would become an essential component of modern bioinformatics systems for medical applications. NJIT Faculty: Jason Wang

A Family-Based Framework for Quality Assurance of Biomedical Ontologies Quality assurance of biomedical ontologies deals with medical terminologies and ontologies, a subject of study that is a sub-field of Medical Informatics. Many biomedical terminologies are measured in the tens of thousands to hundreds of thousands of terms, including drug names and their chemical ingredients, symptoms, diagnoses, body parts, medical procedures, medical devices, infectious agents, accidents, etc. Unfortunately, these medical terminologies are hard to understand and contain inconsistencies. Understanding them and finding inconsistencies with textual representations is close to impossible, therefore graphical representations are used. In a graphical representation, biomedical terminologies appear as networks where the terms are symbolized as boxes and the relationships between pairs of terms are symbolized as arrows. Any naïve approach to visualizing these networks on a computer screen fails. The core elements of this research are therefore (1) developing small abstraction networks that summarize large biomedical terminologies; (2) visualizing abstraction networks on a computer screen in a manner that is easier to comprehend than the original terminologies; and (3) performing quality assurance on the original terminologies by using the abstraction networks to find and remove inconsistencies. As biomedical terminologies are increasingly used in applications such as electronic health records (EHRs), ensuring that terminologies are free of inconsistencies helps ensure the correctness of these applications. NJIT Faculty: James Geller, Yehoshua Perl, Michael Halper, Chris Ochs
Funding: National Institutes of Health (NIH) – National Cancer Institute

A Cyber-Informatics Approach to Studying Migration and Environmental Cancer Risk The World-Wide Web (Web 1.0) and online social media (Web 2.0) have revolutionized the ways medical knowledge is disseminated and health information is exchanged and shared among patients, supporters, and health care providers. Online patient communities have been expanding at an impressive rate with millions of active participants from all age groups. Recent studies on researching and analyzing social media contents for health-related applications show that this increasing cyber-trend leads to valuable knowledge, traditionally acquired with scientific methods such as observational epidemiological studies. This new mode for information acquisition is particularly advantageous for studies requiring long periods of data curation. We leverage the power of online contents, including user-generated contents on social network sites, to tackle the National Cancer Institute's second provocative question on complex migration patterns and their effect on environmental cancer risk. We hypothesize that the rich amount of personal information shared openly among cancer patients and cancer-free people online can be effectively mined to generate new knowledge on the topic, which cannot be easily uncovered with conventional migrant studies in our modern economy with population mobility patterns far more complex and dynamic than those observed in the past. To achieve our goal, we are building upon our unique cyber-informatics experience at the Oak Ridge National Laboratory (ORNL) on ultra-scale searching, identifying, and understanding free-structured web content. This study has the potential to provide a powerful complementary approach to the standard paradigm of observational epidemiological research. It will offer a fully automated and cost-effective way to discover new trends and monitor evolving ones on the impact of modern population migration patterns and environmental cancer risk. Such information could help cancer epidemiologists and health policy makers generate and prioritize study hypotheses worth testing with carefully controlled and properly powered (but also long-term and costly) epidemiological studies.
NJIT Faculty: Songhua Xu Funding: National Institutes of Health (NIH) – National Cancer Institute (NCI)

Statistical Modeling for Genomic Data Analysis We investigate the problems in genetics and genomics as interesting statistical problems and develop novel statistical models and computational methods to solve these problems. We are particularly interested in developing statistical methods for analysis of high-dimensional genomic data. Our recent methodological contributions include change-point models for sequencing data analysis, machine-learning approaches for disease risk prediction, and graphic models for pattern recognition and integrating multiple-source data. These works have been published in prestigious journals including *JASA, Biometrika, AJHG, AOAS, PLoS Genetics*, and *Bioinformatics*. NJIT Faculty: Zhi Wei Funding: National Institutes of Health (NIH)

0/1 Loss-Based Algorithms for Classification in Machine Learning Classification based on minimizing the number of classifications, also known as 0/1 loss, suffer less from misclassified outliers compared to mainstream convex approximations. At the same time, optimizing the 0/1 loss is NP-hard and has multiple local minima. We have developed and studied new local search methods and a single-layer network model for non-linear classification based on 0/1 loss. We study our methods empirically on real data and achieve lower mean error than state-of-the-art methods. We are also studying image- and multi-class classification and extending our algorithmic approaches to regression and unsupervised learning. NJIT Faculty: Usman Roshan

HUMAN-CENTRIC COMPUTING

Social Media and Emergency Response Increasingly, the public is using social media such as Twitter and Facebook during disasters to share information. This information is generally very timely (e.g., the first Tweets about an earthquake occur less than a minute after onset) and potentially valuable for providing situational awareness to help professional emergency managers (EMs). However, many agencies do not make use of this potentially valuable resource, except for posting announcements for the public. A series of three studies has been conducted exploring perceived barriers to the use of social media among EMs. The first was based on semi-structured interviews in order to make sure that we had identified all of the major perceived barriers in order to be able to construct a comprehensive set of surveys. The second was an online survey completed by over 200 county-level emergency managers in the United States, which measured how and to what extent they are currently using social media, and explored a range of barriers to use, including organizational policies and procedures, lack of trained staff, and technology-related issues such as lack of perceived trustworthiness and information overload. The third study, an online survey completed by almost 500 county-level emergency managers in the United States, focuses on their reactions to a number of software tools that could help solve the technical problems. Each of these studies was presented as a paper at an annual meeting of ISCRAM (Information Systems for Crisis Response and Management), and journal-length articles are under consideration. Thus far the project has been unfunded but we are currently exploring funding for follow-up workshops to devise design guidelines for software to filter and organize social media posts related to disasters. NJIT Faculty: Starr Roxanne Hiltz Funding: Travel grant from the University of Agder, Norway

Modeling Disaster Response and Infrastructure Interdependencies A series of studies with colleague Jose Gonzalez at the University of Agder, Norway; Victor Banuls Silvera, University Pablo de Olivades, Spain and a former Fulbright scholar in the Information Sciences Department at NJIT; and others, was begun during the visit at University Carlos III in Spain, and has received some travel-funding support from the University of Agder. These studies use either systems dynamics modeling or a combination of online Delphi studies and Interpretive Structural Modeling. Typically they involve a small set of experts (ranging from three to about 20 in various papers) who make the estimates of relationships that are then mathematically combined into the model.
NJIT Faculty: Starr Roxanne Hiltz, Murray Turoff

Distributed Group Support Systems These systems use decision support software (such as voting with feedback) and structured communication to support groups and virtual teams who are located in different places and work together through computer-mediated communication. A series of five multi-year grants awarded to NJIT faculty by the National Science Foundation between 1978 and 2008 supported many Ph.D. students involved in this research. Although this research has not been funded by new grants since 2010, the principals continue to collaborate with former doctoral students who are still involved in research in this area. Much of this research, carried out at the former Collaborative Systems Laboratory at NJIT, was funded by the UPS Foundation, the National Science Foundation and the state of New Jersey. NJIT Faculty: Starr Roxanne Hiltz, Jerry Fjermestad, Murray Turoff

Context-Aware Mobile Social Matching A series of laboratory and field studies were performed of how location-aware mobile social software might help connect students on campus and expand their friendship networks.
NJIT Faculty: Quentin Jones, Starr Roxanne Hiltz Funding: National Science Foundation (NSF) Cyber-Human Systems (CHS)

Motivations and Behaviors of Daily Fantasy Sports This project examines the psychological motivations and habits of people who play online Daily Fantasy Sports (DFS). Unlike those who engage in sports betting, DFS players conduct research and engage in multiple information-seeking and information-processing behaviors. We aim to understand cognitive models of how people make both information and financial decisions through both qualitative and quantitative methodologies. NJIT Faculty: D. Yvette Wohn Funding: Yahoo

The Influences of Social Media on Society Ten different publications in the past year have examined different aspects of how social media influence people's attitudes and behaviors. Ongoing research includes news consumption via social media, habits, and linkages between social media use and well-being. NJIT Faculty: D. Yvette Wohn

The Impact of Undetectable, Wearable Computing Devices on Classroom Dynamics Wearable computing devices, such as smart glasses, essentially will become invisible in the near future. As students inevitably bring their wearable technology to class and use them undetectably, continuously online (UCO), how will student and instructor behavior evolve regarding learning, teaching, assessment and classroom dynamics? Can behavior, motivation, learning, teaching, and assessment be guided and improved so we gain the maximum benefit for learning and teaching? Can we design tools to help students become aware of and learn to self-regulate their UCO activities, minimize distraction and negative behavior, and become responsible digital citizens? NJIT Faculty: Michael Bieber, Starr Roxanne Hiltz, D. Yvette Wohn, Songhua Xu

INFORMATION INTEGRATION AND INFORMATICS

Information Filtering by Multiple Examples (IFME) A key to successful online searching lies in how users express keywords as queries. However, many users experience the difficulty of finding the right keywords, especially when the information need is complex. This project proposes a framework named Information Filtering by Multiple Examples (IFME) to solve this issue by using relevant sample documents that a user provides to search for similar results without the need for the user to compose a query. This framework explores the following issues: (1) solving the topic diversity problem in predicting or modeling the user's information need derived from user's samples; (2) tackling the class imbalance problem (too few positive instances and too many unlabeled instances) using under-sampling; (3) identifying the most effective positive-unlabeled (PU) learning algorithms for IFME; (4) adopting ensemble learning techniques to improve the effectiveness for IFME further; and (5) efficiently reducing feature dimensions by selecting the most important document features. The experimental results show that the proposed ensemble learning approach augmented with under-sampling outperformed baselines. The study also finds that when users provide 10 to 20 sample-relevant documents, the system achieves good results. NJIT Faculty: Brook Wu

Task-Based User Profiling for Query Refinement (TOQUE) A user's search task, consisting of a sequence of search queries serving the same information need, can be treated as an atomic unit for modeling user search preferences. This study proposes to personalize query refinement by using task information extracted from user logs. The framework starts by identifying a user's search tasks: a cross-session-based method proposed to discover tasks by modeling the best-link structure of queries, based on the commonly shared clicked results. Then, it represents the user's dynamic task-level search interests with a multi-descriptor model to represent long- and short-term user interests extracted from search tasks and sessions. This multi-descriptor model represents the user's interests at the task level which is used to re-rank candidate queries for personalization through modules of task identification and update. In the selective personalization process, TOUQE utilizes a two-step algorithm to improve the ranking of candidate queries by assessing the task dependency via exploiting a latent task space. If the task dependency is not high enough, there will be no personalization and candidate queries will be presented to the users in their original order. Experimental results show that the proposed framework resulted in significantly more relevant candidate queries and thus shortened the search sessions. NJIT Faculty: Brook Wu

THEORY AND OPERATIONS RESEARCH

Efficient Monte Carlo Methods for Characterization of Safety Margins of Nuclear Power Plants A recent international effort of the Nuclear Energy Agency formulated a new risk-analysis framework, known as the risk-informed safety-margin characterization, to address several critical issues facing the nuclear industry today. The impact of these developments, which include extending operating licenses for aging facilities and running them at higher power levels, can profoundly affect safety margins that were previously deemed acceptable, and there is an urgent need to better understand the risks in the changes. Indeed, the catastrophic 2011 accident at the Fukushima Daiichi plant in Japan clearly shows the utmost importance of accurately measuring risk. This project includes research on novel approaches to dramatically reduce statistical error in Monte Carlo simulations to evaluate risk. In addition to performing efficient nuclear safety analyses, the techniques may also be applied to attack problems in civil and mechanical engineering, catastrophe modeling, and finance. The research combines theoretical and practical components, including designing new Monte Carlo algorithms, formally establishing their validity, and implementing them to analyze risks of existing nuclear plants. NJIT Faculty: Marvin K. Nakayama Funding: National Science Foundation (NSF)

Efficient Monte Carlo Methods for Estimating Risk The project focuses on the development of efficient simulation methods for estimating risk measures and other performance metrics. Variance-reduction techniques (VRTs) for estimating a quantile and constructing confidence intervals for it are key focuses of the research. Quantiles are commonly employed to measure risk in a diverse range of application areas, such as finance and nuclear engineering. For example, the Nuclear Regulatory Commission (NRC) regulations sometimes specify that risk is measured using a "95/95 criterion," which requires computing a 95 percent confidence interval for a 0.95-quantile. To satisfy the NRC requirements, practitioners currently resort to crude Monte Carlo, which can be hugely inefficient, so the research on constructing confidence intervals for quantiles with VRTs has the potential for altering the way uncertainty and safety analyses are performed in the nuclear industry. In finance, where a quantile is known as a "value-at-risk," there are banking rules for capital requirements given as a function of a 0.99-quantile. Other fields in which the research on efficient quantile estimation can be applied include project planning, telecommunications, manufacturing, supply chains, and computing. NJIT Faculty: Marvin K. Nakayama Funding: National Science Foundation (NSF)

EXTREEMS-QED: Research and Training in Computational and Data-enabled Science and Engineering for Undergraduates in the Mathematical Sciences The program immerses NJIT undergraduates in courses and group projects relating to mathematical computation and data analysis, and promotes faculty professional development through participation in workshops on computational science education. Students are exposed to computational issues in multidisciplinary situations, including in industrial settings. The program includes the development of new

courses and enhancements of existing ones that will permanently enrich the Computational and Data Enabled Science and Engineering (CDS&E) curriculum at NJIT. Some of the enhancements are in courses that are taken by students in a wide variety of majors, and the changes will affect their careers in science and technology. The group research projects use and develop fast and accurate algorithms to study problems in medicine, fluid dynamics, acoustics, climate modeling, and risk assessment. Many of the projects will lead to improvements in computational techniques for data analysis, such as dimension reduction, data assimilation, and Monte Carlo methods. Valuable information will be gained and disseminated on effective ways to educate students in these computational techniques. NJIT Faculty: Marvin K. Nakayama, Michael Siegel, Zoi-Heleni Michalopoulou, David Horntrop, Ji Meng Loh Funding: National Science Foundation (NSF): Division of Advanced Cyberinfrastructure and Division of Mathematical Sciences

Multi-Core, Many-Core, and Parallel Algorithm Design, Analysis, and Implementation Models of computation that abstract and also capture parallelism in the presence of multiple memory hierarchies and multiple cores are being studied and developed. Multi-core systems are limited parallelism architectures with fast inter-core communication but limited fast memory availability. Abstracting the programming requirements of such architectures in a useful and usable manner is being explored. NJIT Faculty: Alex Gerbessiotis

MULTI-MEDIA, TREE AND GRAPH DATABASES, AND THE SEMANTIC WEB

Mining and Summarizing Homomorphic Patterns from Large Trees and Graphs Extracting frequent patterns which are hidden in trees and graphs is central for analyzing data and is a basic step for other data-mining processes, including association rule mining, clustering, and classification. Many efficient pattern mining algorithms have been discovered in the last two decades, yet most of them do not scale to the type of data presented to us by Big Data applications. In this project, we design novel algorithms for extracting homomorphic tree and generalized DAG patterns (which summarize multiple patterns) from large trees and graphs. Our approach exploits methods for encoding tree data and well-studied modern holistic twig-join algorithms for evaluating tree- and graph-pattern queries over tree data. It leverages recent database results on materializing views as compressed bitmaps and on answering queries using bitmap views in order to efficiently compute in an incremental way the support of candidate frequent patterns from previously identified ones. NJIT Faculty: Dimitri Theodoratos

Searching the Semantic Web with Keyword Queries The resource description framework (RDF) provides a standardized representation of linked data, which is the backbone of the Semantic Web. Nowadays, RDF repositories are growing constantly. Keyword search is a popular technique for querying and retrieving information from the Web. Disambiguating the user's intention in posing a keyword query over an RDF graph database and efficiently retrieving relevant results among a plethora of candidates are challenging problems in RDF keyword search. To address these problems, in this project, we follow an approximate schema-aware approach that exploits a structural summary of the RDF graph to extract pattern graphs for keyword queries. The pattern graphs cluster together results with the same properties. In this context, we devise techniques for dealing with the following issues: (a) designing a clustering hierarchy which allows the user to efficiently select the relevant results without computing entirely the result set; (b) relaxing patterns graphs in order to enable the extraction of additional results potentially of interest to the user; and (c) balancing relevance and diversification in the pattern graph sets for keyword queries in order to maximize the satisfaction of the average user. NJIT Faculty: Dimitri Theodoratos

Multimedia Retrieval and Data Visualization This research has introduced innovation in developing a web-based framework to help small businesses in New Jersey find collaborative opportunities with the Department of Defense. The information is clearly defined by searchable criteria such as type of business, type of products and/or services, regions, and type of industry (such as computer or pharmaceutical). It is designed to eventually provide a national model to increase and improve relationships between the Department of Defense and the small business workforce. NJIT Faculty: Chengjun Liu, Yi Chen, Zhi Wei, Taro Narahara Funding: Department of Defense

RESEARCH ON COMPUTING EDUCATION

Computing Education In the United States alone, jobs in computer science are increasing at double the national average and are among the top paying fields, but there are not enough people trained to fill these roles, especially in programming. It is unclear who is attracted by online educational resources, and how they affect engagement and learning. This project creates technologies to address these issues., including the freely available game, Gidget (helpgidget.org), which has reached thousands of people all over the world. It has shown to be engaging, has demonstrated measurable learning outcomes, and has been successful at reaching underrepresented groups (45 percent of users are female). NJIT Faculty: Michael J. Lee

Integrating Learning Resources for Information Security Research and Education (iSECURE) This project links and integrates multimedia teaching materials (slides, videos and textbooks) used in Information Security courses based on a security ontology. The semantic linking of the multimedia course materials allows students to search and compose multimedia and interactive course materials based on content, medium, and learning styles, enabling flexible personalized learning. To achieve these goals, the project is focusing on several research tasks: (1) segmenting and annotating learning media based on their learning content; (2) building a security ontology to be used for annotating and querying of the learning objects; (3) adapting lecture content to student learning styles; and (4) evaluating the usability and effectiveness of the linked multimedia security learning system. iSECURE is a collaborative project with Montclair State University (Edina Renfro-Michel) and CUNY Staten Island (Soon Ae Chun). NJIT Faculty: Vincent Oria, Reza Curtmola, Jim Geller Funding: National Science Foundation (NSF)

PROFILES OF RESEARCHERS

MICHAEL BIEBER, professor and associate chair, Information Systems Department, conducts research in the learning sciences, cyberlearning, lightweight systems integration, relationship analysis (as part of the software engineering process), automatic link and meta-information generation, and hypermedia. Ph.D., University of Pennsylvania.

CRISTIAN BORCEA, associate professor and chair, Computer Science Department, has more than 15 years of experience in the fields of mobile computing and sensing; ad hoc and vehicular networks; and cloud and distributed systems. Specifically, he focuses on designing, implementing, and evaluating system architectures, network protocols, middleware, programming models, and security mechanisms for networked systems. He has published more than 70 articles in journals and conferences proceedings, and he is a co-author of *Mobile Crowd Sensing*, which will available in August 2016. Ph.D., Rutgers University.

REZA CURTMOLA, associate professor, Computer Science Department, is the recipient of the 2011 NSF CAREER Award for research in securing cloud services. His other research interests include applied cryptography and security aspects of the software supply chain, mobile computing, crowdsensing, and wireless networks. He spearheaded the creation of the M.S. degree program in Cyber Security and Privacy at NJIT and he led the effort to certify the university as a Center of Academic Excellence in Cyber Defense Education by the NSA/DHS. Ph.D., Johns Hopkins University.

XIAONING DING, assistant professor, Computer Science Department, came to NJIT from Intel Labs at Pittsburgh (now Intel ISTC for Cloud Computing), where he was as a computing innovation fellow. His current research focuses on building highly efficient cloud computing infrastructures for new application domains, such as mobile computing applications and Big Data applications, on top of new computing hardware architectures, such as multicore and non-volatile memory. His research areas include parallel and distributed systems, operating systems, database systems, and computer architecture. Ph.D., Ohio State University.

NARAIN GEHANI, professor, Computer Science Department, was previously the chair of the department and dean of the Ying Wu College of Computing Sciences. After 23 years at Bell Labs (Lucent Technologies & AT&T), where he rose to become vice president of the Communication Software Research Lab, he became president of Maps On Us (a Lucent commercial website), later SwitchBoard. He has created several software systems, including the Ode database and Concurrent C/C++ parallel programming language. He holds several patents and is the author of many books and papers in computer science. Ph.D., Cornell University.

JAMES GELLER, professor and former chair, Computer Science Department, is associate dean for research of the Ying Wu College of Computing Sciences and co-director of the Structural Analysis of Biomedical Ontologies Center (SABOC). He is a fellow of the American College of Medical Informatics and a recipient of the NJIT Master Teacher Award. His research interests include the use of ontologies in medical informatics and in education. His recent research has concentrated on the summarization, visualization, and auditing of biomedical terminologies and ontologies. He has published more than 175 articles in journals and conference proceedings. Ph.D., State University of New York at Buffalo.

ALEX GERBESSIOTIS, associate professor, Computer Science Department, has research interests that include multi-core and parallel computing, design, analysis, and implementation of sequential and also parallel latency-tolerant algorithms, combinatorics, and random graphs. Ph.D., Harvard University

MICHAEL HALPER is professor, Information Systems Department, and director of the Information Technology Program. His research interests include conceptual and object-oriented data modeling, extensible data models, and medical informatics with a focus on terminologies and ontologies. His recent work on terminologies and ontologies has dealt with issues such as: partitioning and the derivation of various abstraction networks (e.g., schemas, taxonomies, and meta-schemas); enhanced representation and modeling; and quality-assurance (QA) methodologies.

STARR ROXANNE HILTZ, distinguished professor emerita, Department of Information Systems, has held visiting research and/or teaching appointments at the University of Salzburg, University Carlos III of Madrid, the University of Agder in Norway, and the Tecnun, University of Navarro in Spain. Her research interests include virtual teams and online communities, online learning, emergency response information systems, and social media. The author of numerous books, she is the recipient of many awards, including the 2012 Making a Difference Award from the Association for Computing Machinery, given annually to an individual nationally recognized for work related to the interaction of computers and society, and recognized for her "research and teaching on the issues of ubiquitous network computing and online learning, and their social impact." Ph.D., Columbia University.

MICHAEL J. LEE, assistant professor, Information Systems Department, specializes in the areas of computing education research (CER) and human-computer interaction (HCI). He focuses on ways to engage and instruct people in programming and to measure their progress. He created an online game (Gidget) designed to teach people of all ages how to program by solving debugging puzzles and developed a tool for children with a programmable robotic dinosaur. Lee is also the recipient of a National Science Foundation (NSF) grant to work with international collaborators to examine the effectiveness of Gidget in different cultures. Ph.D., University of Washington.

CHENGJUN LIU, professor, Computer Science Department, conducts research in pattern recognition, machine learning, computer vision, image and video processing, and biometrics. He has developed advanced pattern recognition and machine-learning methods, new color models, and new similarity measures, and applied them to address large-scale and grand-challenge problems. His research has been published in the major journals and conferences in his field, such as the IEEE *Transactions* journals and the IEEE conferences. Ph.D., George Mason University.

ALI MILI, professor, Computer Science Department, holds broad interests in software engineering, ranging from the technical to the organizational to the managerial aspects of the field. He has co-authored 11 books, 12 book chapters and 230 articles in journals and conference proceedings. He has worked in Algeria, Australia, Austria, Canada, China, France, Japan, and Tunisia. Ph.D., University of Illinois; Doctorat ès Sciences d'Etat from the University of Grenoble.

MARVIN K. NAKAYAMA, professor, Department of Computer Science, served as the area editor for the Simulation Area of the INFORMS *Journal on Computing* from 2007-2016, and has been on the editorial board of the ACM *Transactions on Modeling and Computer Simulation* since 1998. His research interests include stochastic simulation, modeling, applied probability, statistics, energy, and analyzing risk. His papers have appeared in journals such as *Operations Research, Management Science, Mathematics of Operations Research, Annals of Statistics, Advances in Applied Probability, IEEE Transactions on Computers*, and the *ACM Transactions on Modeling and Computer Simulation*. He is the recipient of a CAREER Award from the National Science Foundation. Ph.D., Stanford University.

IULIAN NEAMTIU, associate professor, Department of Computer Science, is a recipient of the CAREER Award from the National Science Foundation, and the University of California, Riverside, Regents' Fellowship award, as well as two Google Research Awards. His research areas span programming languages, security, software engineering, and smartphones, with an overarching goal of making software and smartphones more secure, efficient, and dependable, as well as easy to maintain and modify. Ph.D., University of Maryland, College Park.

CHRISTOPHER OCHS, postdoctoral research associate, Computer Science Department, focuses on algorithms for creating structural summaries of ontologies, ontology quality assurance methodologies, and software tools for ontology summarization and quality assurance. Ph.D., NJIT.

VINCENT ORIA, associate professor, Computer Science Department, is the recipient of the 2014 NJIT College of Computing Sciences Outstanding Achievement in Research Award and the 2015 ACM (Association for Computing Machinery) SIGMOD (Special Interest Group on Management of Data) Test of Time Award. His research interests include spatio-temporal databases, multimedia databases and search in high-dimensional spaces. Ph.D., Ecole Nationale Supérieure des Télécommunications (Telecom-Paritech), Paris, France.

YEHOSHUA PERL, professor, Computer Science Department, is co-director of the Structural Analysis of Biomedical Ontologies Center (SABOC) research center and an ACMI (American College of Medical Informatics) Fellow. He has published extensively in the fields of algorithms, object-oriented databases, and biomedical terminologies. His recent research concentrates on the summarization, visualization, and quality assurance of biomedical terminologies and ontologies. Ph.D., Weizmann Institute, Israel.

KURT ROHLOFF, associate professor, Computer Science Department, initiates and executes advanced research and development projects for cyber-security and secure computing at scale. His research interests include secure cloud computing, fully homomorphic encryption (FHE), practical encryption, high performance distributed computing, Big Data, graph data analytics and rare-event modeling. He is currently leading the development of a best-of-breed open source library for post-quantum and homomorphic encryption. Ph.D., University of Michigan.

USMAN ROSHAN, associate professor, Computer Science Department, is director of the Bioinformatics Program. His research interests in bioinformatics include: sequence alignment, evolutionary tree reconstruction, genome-wide association studies, and disease risk prediction; and in computer science: data science, machine learning, high-performance computing, and GPU algorithms. Ph.D., University of Texas at Austin.

ANDREW SOHN, associate professor, Computer Science Department, founded a venture capital-backed software company in Silicon Valley in 1999 that marketed a dynamically scalable computing infrastructure, now known as Cloud computing. Previously he was a National Science Foundation fellow at the Center of Excellence in Computer Science at the Electrotechnical Laboratory of Ministry of Economy, Trade and Industry (METI), in Japan; a NASA Joint Venture fellow at NAS (Numerical Aerospace Simulations Division) Parallel Benchmarking of NASA Ames; and an invited scientist at NASA Research Institute for Advanced Computer Science and Berkeley National Laboratory. Ph.D., University of Southern California.

MURRAY TUROFF, distinguished professor emeritus, was a founder of Information Systems for Crisis Response And Management (ISCRAM) in 2004 and served on the board of directors until 2014. He has authored or co-authored eight books including the award-winning *The Network Nation: Human Communication via Computer* (with his wife Starr Roxanne Hiltz), which became the defining document that predicted all the current web based communication systems in use today. He has been chair of both the Information Systems Department and the Computer Science Department at NJIT, as well as serving simultaneously as a faculty member of the Rutgers Graduate School of Management. He is currently active in research on the use of information systems for crisis management. Ph.D., Brandeis University.

DIMITRI THEODORATOS, associate professor, Computer Science Department, is the recipient of an IBM faculty award and an excellence in teaching award in graduate instruction from NJIT. His research interests include database theory, data integration, data warehousing, query processing and optimization, and pattern mining. His research currently focuses on keyword search over tree data and graph data (and in particular the Semantic Web), and on extracting homomorphic patterns from trees and graphs. He has published more than 75 papers in journals, books, and conference and workshop proceedings. Ph.D., Laberatoire de Recherche en Informatique, Université Paris-Sud, France.

GUILING "GRACE" WANG, associate professor, Computer Science Department, is co-author of a number of papers in journals, conference proceedings and books. Her research interests include mobile computing, intelligent on-road transportation, and wireless sensor networks. Ph.D., Pennsylvania State University.

JASON T. L. WANG, professor, Computer Science Department, is director of NJIT's Data and Knowledge Engineering Laboratory. The executive editor of the World Scientific Book Series on Science, Engineering and Biology Informatics, he has published more than 150 papers and eight books, the most recent of which is *Computational Intelligence and Pattern Analysis in Biological Informatics*. His research focuses on data mining, machine learning and computational biology. He is a founding chair of the ACM Special Interest Group on Knowledge Discovery and Data Mining Workshop on Data Mining in Bioinformatics. Ph.D., Courant Institute of Mathematical Sciences, New York University.

ZHI WEI, associate professor, Computer Science Department, is an editorial board member of *PLoS ONE*, *Frontiers in Bioinformatics and Computational Biology*, and *Frontiers in Applied Genetic Epidemiology*. His research interests include statistical modeling, machine learning, and Big Data analytics, with applications to many data-enriched fields such as bioinformatics, genetics, social media, web analytics and advertising. He has published in various journals and conference proceedings, including *JASA, Journal of the American Statistical Association; Biometrika; Annals of Applied Statistics, AJHG, Nature, Nature Medicine*, and *Proceedings of the National Academy of Sciences*. Ph.D., University of Pennsylvania.

D. YVETTE WOHN, assistant professor, Information Systems Department, is interested in understanding long-term use of technology through the lens of "soft sustainability," which focuses on the human social factors of sustainable technology use. She is particularly interested in understanding social systems for education and health purposes and is developing social mobile apps to help alleviate loneliness and facilitate social support. She is the recipient of a 2015 Yahoo Faculty Research and Engagement Program Award. Ph.D., Michigan State University.

BROOK WU, associate professor and chair of the Information Systems Department, collects and analyzes user search and browsing behavior data to improve user-interest representation, personalization of search queries and search results, and user satisfaction in fulfilling information needs. Her teaching and research interests include text mining, user search behavior, and information retrieval. She is the co-editor, with Min Song, of the *Handbook of Research on Text and Web Mining Technologies*, and co-author of numerous papers in professional journals. Ph.D., State University of New York at Albany.

CHASE QISHI WU, associate professor, Department of Computer Science, is also on the research staff of Oak Ridge National Laboratory, where he works on a number of high-performance networking projects and Big Data computational science projects. His research interests include Big Data, high-performance networking, parallel and distributed computing, sensor networks, scientific visualization, and cyber security. His research in computing develops high-performance workflow solutions to manage the execution and optimize the performance of large-scale scientific workflows in heterogeneous network environments. He has published more than 170 research articles in conference proceedings, journals, and books. Ph.D., Louisiana State University.

SONGHUA XU, assistant professor, Information Systems Department, and former Wigner Fellow at Oak Ridge National Laboratory, has 15 approved patents. His primary research interests include information retrieval and management, web search and data mining, innovative applications of artificial intelligence, and intelligent systems for biomedical applications. He is also interested in visual computing, multimedia, human-computer interaction, computer graphics and visualization, as well as digital arts and design. Ph.D., Yale University.