Faculty members from many different departments identify their research as part of a system, whether they are NEMS (nano-electromechanical systems), MEMS (microelectromechanical systems), information systems or complex management systems. Many are working at the edge of CyberBioPhysical Systems™, the area where the biological, physical, and digital worlds fuse and where advances will be made that will impact key challenges in health, sustainability, information and infrastructure.
Ask Professor Steven Bellovin about computer privacy and he might start by discussing aviation. “The technology is so good, there are no single causes of airplane crashes any more. But when complicated systems interact in complicated ways, you have unexpected failures,” he said.

Bellovin has seen that complexity emerge on the Internet. Thirty years ago, he helped create USENET, a precursor of today’s Internet forums. He wrote the first book on Internet security, and is now creating software to simplify network security. He remains an important voice in public discussions about privacy.

“Computers interact with the world around them,” Bellovin said. “We cannot be only scientists or engineers. We have to bring our knowledge to the debate. We have no more right to a policy opinion than anyone else, but no less right either.”

He sees the Internet’s interconnected technologies eroding personal privacy. For example, nearly all commercial Websites collect information about users. While some keep that information private, others do not. Anyone can crosscheck-for-sale databases to unearth personal information.

“In 1994, Congress mandated that telecommunications switches include technology to make it easier to tap phones. We could tell this would be abused. Sure enough, someone tapped one hundred people in Greece, including the prime minister. When we see proposals like this, it is our obligation as specialists to say something.”

Some privacy mechanisms fail because large Websites actually consist of many different services. Not all of them share the same privacy policies. Facebook, for example, stored pictures on servers that did not enforce privacy rules. Hackers could scrape supposedly private data by entering through those servers.

A third area of concern is anonymization, a process that wipes identifying data from database records. Yet many companies can use anonymized data to build detailed records of individuals. Google, for example, captures queries, offers check-out services that record purchases, and owns Double-Click, which tracks clicks for advertisers.

This could enable it to create detailed profiles. “Some people want to see ads about things they like. Others find it creepy that somewhere there’s a repository of all your information,” Bellovin said.

“Part of the solution is educational,” he continued. “We can teach people to protect their privacy. But it’s also a technology issue.”

His group is looking at better ways to preserve privacy. This includes creating unlinkable aliases, improving the privacy of database searches, and encrypting advertising clicks so merchants cannot access private information.

B.A., Columbia, 1972; M.S., North Carolina (Chapel Hill), 1977; Ph.D., 1982
Gridlock doesn’t just happen on highways. Interlocking congestion that prevents movement is also a threat to the Internet. As more people exchange more information on a more frequent basis, the Internet’s traffic management system (routers) is forced to use more energy to forward and receive data between computer networks. As routers lose ground against traffic demands, performance bottlenecks occur.

Photonic, the science and technology of generating and controlling photons, could ease up electronic traffic jams by providing the solution to internet gridlock. Through photonics, the potential exists to achieve advanced information traffic management performance along with energy efficiency by symbiotically merging the computation-communications infrastructure. Optical routers would transmit data as light, avoiding unnecessary electronic processing. In addition, they would use less power consumption while manipulating gargantuan amounts of data with complete format transparency in a smaller device footprint.

Keren Bergman leads the Lightwave Research Laboratory at the University. She investigates the realization of dynamic optical data routing in transparent optical interconnection networks. Through this work, she is developing potentially disruptive technology solutions with ultra-high throughput, minimal access latencies, and low power dissipation that remain independent of data capacity. These solutions will ultimately capitalize on the enormous bandwidth advantage enabled by dense wavelength division multiplexing.

Her work on large scale optical networks focuses on embedding real-time substrate measurements for cross-layer communications. As envisioned by the community, this suite will support a wide range of network science and engineering experiments such as new protocols and data dissemination techniques running over a substantial fiber optic infrastructure with next-generation optical switches, novel high-speed routers, city-wide experimental urban radio networks, high-end computational clusters, and sensor grids.

Bergman’s research in large-scale optical switching fabrics includes cross layer optimized optical substrate and embedded real-time measurements. Her work in optical interconnection networks for high-performance computing systems includes data vortex optical packet switching fabric, optical network interface card and scalable optical packet buffers. Her work in integrable interconnection network systems and subsystems includes parametric optical processes and systems and nanophotonic optical broadband switches. Her work in inter- and intra-chip multi-processor interconnection networks includes on- and off-chip photonics communications for multi-processor systems and silicon photonic devices for networks-on-chip.

B.S., Bucknell University, 1988; M.S., MIT, 1991, Ph.D., 1994
In early 2008, few investors saw the whirlwind coming. The financial crisis was what economists call a black swan, an earthshaking event so unlikely, no one anticipates or plans for it. Assistant Professor Jose Blanchet would like to rectify this situation.

“I study Black Swan events by using probabilistic methods. That doesn’t mean I predict them. Instead, I use computers to understand how they evolve,” Blanchet said. His goal is to help investors see the warning signs of extreme events before they occur, while they still have time to respond.

Blanchet does this by building realistic computer models of portfolios. As they evolve, he shocks them with random events, such as bond defaults and bankruptcies. Ordinarily, the portfolio absorbs the hits. Rarely, very rarely, a combination of random shocks sends values crashing, just as cascading events did to real portfolios in 2008.

The shocks, Blanchet explained, must be truly random. “If you try to model a crisis that simulates these events, you could get it wrong and it will not reflect reality,” he said.

“For example, suppose this is 2006 and you want to see what happens if lots of people default on their mortgages. Rather than start with a bankruptcy, we want to start with the events that cascade to create the bankruptcy. We let the probability model capture the events that occur naturally, even if they are rare.”

“We look at extreme events in such contexts as queueing networks and risk management of financial and insurance portfolios. We want to understand what happens when there are huge backlogs or when companies post enormous losses. What are the consequences of that? What is the likelihood?”

Ordinarily, it would take a week or two to run enough simulations to generate a single black swan. That is far too slow to build a large enough database to study these events for similarities and differences.

To get around this problem, Blanchet devised algorithms that generate black swans rapidly. He then runs hundreds of simulations using a variety of portfolio models to see how they behave.

“We have a family of models that capture the features we want to study, and a computational tool that lets us observe these events as they unfold,” he said.

“It’s like watching a crack in a dam,” Blanchet said. “Most of the time, nothing happens. But sometimes it propagates and then the dam goes.”

B.S., Instituto Tecnológico Autónomo de México, 2000; Ph.D., Stanford, 2004
As microprocessors grow more powerful and complex, engineers dream of putting an entire computer system on a single chip. Such chips would be smaller, faster, and more energy efficient than today’s designs. To get there, though, we will need to reinvent how we design chips, Associate Professor Luca Carloni argues.

Today, engineers create microprocessors using tools that help them build circuits from libraries of proven designs. Yet new technologies pose many problems for traditional tools. In the past, for example, chips synchronized all operations with a single clock. “Compared to the times needed for computation, on-chip communication was basically instantaneous,” Carloni explained. “Today, local calculations run so fast, it takes several clock cycles for remote signals to arrive. This is a physical issue we need to address.”

Carloni ticks off other problems. New chips have multiple processors, or cores, whose parallel operations create new challenges in programmability. Billions of transistors create new levels of complexity and generate lots of heat that is hard to remove. Resolving these issues has extended the amount of time and design iterations needed to create new chips.

Those same emerging technologies also offer new opportunities. Instead of trying to develop a system-on-chip with old tools, Carloni proposes reinventing chip architectures and the tools used to design them. “We need to create communication infrastructures that make it easier to integrate new components into our designs,” he said.

His solution is a network on a chip. “Our vision is to create an on-chip communication and control infrastructure. When we have a network that touches every component on a chip, we can dynamically configure the processor to optimize speed or efficiency. We don’t have the solution yet, but we’re working on it.”

He envisions a collection of communications elements—nanoscale wires, switches, and routers—to move data around the chip. The cores would have standard interfaces to plug into the network. A new generation of tools would support component selection and network optimization.

“Instead of designing links between each circuit, you would plug components into a standardized backbone,” Carloni said. “This makes it much easier to design processors. Engineers could continuously upgrade and test new components, then plug them in and know they would work on the chip.”

Networked chips would also support multiple cores running at different clock rates. Chips could assign tasks to different cores to optimize speed or reduce energy use. “The path towards green computing systems starts with more efficient communication infrastructures,” Carloni said.

B.S., Bologna (Italy), 1995; M.S., California (Berkeley), 1997; Ph.D., 2004
In 1987, automated trading programs shoved the market off a precipice. In 2008, a liquidity crisis brought the global financial system to its knees. Associate Professor Rama Cont, who uses probabilistic methods to model financial markets, has studied such system-wide discontinuities for more than a decade. His research on market discontinuities and systemic risk has made him a valued contributor in redesigning financial markets to reduce the impact of major shocks.

“When an epidemic spreads by contact and you cannot vaccinate the whole population, you have to prioritize vaccination resources to prevent further spread. We ask similar questions about market mechanisms that could lead to a financial meltdown.”

Cont takes a system-wide view of financial markets. “We cannot understand why several banks failed simultaneously in 2008 by looking at individual bank portfolios,” he said. “Instead, we must look at the flow of funds and assets in a network of interlinked portfolios. A theoretical physicist by training, he uses the mathematical language of science to analyze financial networks and identify where they are prone to break downs.

In the past, Cont said, regulators promulgated rules that restricted the behavior of individual institutions. “Now, they are trying to look at the market as a whole and assess risks in the entire system. Most markets evolved spontaneously from traders’ needs. Some degree of intervention that strengthens their weakest links can make them less vulnerable to disruption,” he said.

Cont believes clearinghouses can strengthen the system by acting as intermediaries for trades. They would require trading parties to register their transactions. This would increase market transparency about the price – and risk – of derivatives and other instruments that traded at wildly varying prices in the past.

Clearinghouses would also require deposits on all trades. The amounts would rise as institutions take on more risk. The deposits would act as brakes on risk and help compensate for losses if a party defaulted.

Cont is applying his systemic approach to risk management to the design of new derivatives clearinghouses. He is one of the two academics collaborating with the Market Transparency Working Group, a panel of industry officials and regulators charged with redesigning over-the-counter derivatives markets.

“Some people thought that after the market crash, financial engineering was finished. Instead, it raised awareness about the need for rigorous methods for managing risk,” Cont said. “More than ever before, quantitative modeling is in demand now.”

Gautam Dasgupta, Professor of Civil Engineering and Engineering Mechanics, has been working on a wide range of basic engineering problems that span from analytical formulation and mathematical modeling to practical applications. His work in classical civil engineering focuses on analyses of safe but economical design-analysis, such as the dynamic response of nuclear power plants in conjunction with the outwardly radiating waves generated at the soil interface under the action of earthquake excitation, and the effects of acoustic vibrations created by ocean waves on submerged or floating structures, initiating excessive material degradation and damage.

He has also worked with spacecraft engineers, dentists, anthropologists, and historic preservationists. The common thread of his research is to relate the force and deformation of real world systems. His personal interests are reflected in his research applications in computer music and computer-aided graphics.

Dasgupta focuses on research in engineering mechanics. The major objective is to relate the forces (causes) and changes of shapes and sizes (effects) in the real world of uncertainty. Thus his stochastic (models based on probability and statistics) were used by the NASA Glenn Research Center in Cleveland, Ohio, (then NASA Lewis Research Center) in the 1980s to analyze the damage in the main engine turbine blades of the space shuttles. His high-accuracy finite element formulations have been used since the 1990s to analyze the changes in facial bones, including the shape of teeth, as clinicians perform corrective surgery on patients.

Basic engineering mechanics also focus on material properties that change with time under sustained static loading or dynamic impact. The subsequent shape and size changes, which are studied under viscoelasticity and plasticity, have been adapted to computer simulations. Dasgupta derived important theorems for such changes that can vary with frequency and are different in different directions of deforming bodies. In conjunction with his defect-free (finite element) numerical models (of extremely low error), he proposed very large-scale simulations of crash tests for super computers. In order to minimize computing time, a large part of the calculation is carried out algebraically. This led to Dasgupta’s development of symbolic computer programs to derive equations and generate conceptual (elegant) codes for numerical execution in a parallel computing environment.

Dasgupta is working on three-dimensional analyses and computer code generation of incompressible fluids and solid mechanics modeling of complex surfaces and solid objects with zones of mechanical failure to be applied in a wide class of applications ranging from blood flow to the cooling mechanism of wind turbines.

B.Engr. and M.Engr., Calcutta University, Ph.D., University of California, Berkeley
Malaria kills one million victims each year in tropical countries, most of them children. While drugs exist to combat the disease, the malaria parasite develops resistance to these drugs. An effort is now underway to harness a noninvasive electromagnetic-based treatment in the fight against this disease.

The electromagnetic field used in this innovative treatment is the same that is produced in lightning, is responsible for the Northern Lights, and also causes compasses to point in a north-south orientation. It is also found in high-frequency radio waves that bounce from one part of the world to another via antennas to phone networks, TV pictures, and the Internet.

Understanding electromagnetic fields allows scientists to develop smaller and more powerful antennas useful in emergency communication devices and portable radars, or those made flexible with new alloy materials and applied to implantable medical devices. Electromagnetics also has a direct application at the biological level—as in the treatment of parasitic diseases like malaria.

Malaria is caused by a parasite that is transmitted from one human to another by the bite of mosquitoes. In humans, the malaria parasite travels to the liver, where it invades a red blood cell. There, it consumes the cell’s hemoglobin and produces hemozoin, an iron crystal, as a waste product. It then divides into many more daughter parasites that invade other red blood cells. Researchers now understand that the iron crystal remains with the parasite within the host cell. By applying a suitably designed magnetic field, the iron crystals can be made to agitate, rotate, and churn, destroying the parasite before it can multiply further.

Dr. Paul Diament is the lead inventor of the magnetic resonance method of treating the malaria parasite, and is working with biologists at the Columbia Medical School in pursuing this application. He is an eminent researcher in electromagnetics and wave propagation. His teaching and research focus includes microwaves, antennas, optics, radiation statistics, plasmas, wave interactions, relativistic electron beams, and transient electromagnetic phenomena. Along with biomedical applications, his research interests include attempts to make mutual coupling among antennas beneficial rather than detrimental, potentially achieving smaller antenna size.

Dr. Diament is a member of the Institute of Electrical and Electronics Engineers (IEEE), the American Physical Society (APS), the Optical Society of America (OSA), Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.

B.S., Columbia, 1960; M.S., 1961; Ph.D., 1963

Employing Electromagnetics to Treat Malaria

PAUL DIAMENT
Professor of Electrical Engineering
A car takes a curve too fast. Before it spins out, its stability control system kicks in. Its microprocessors calculate and recalculate the right amount of force to apply to each wheel, adjusting the brakes many times per second until the car comes under control. Such critical systems often juggle several events at once. A car’s stability system must calculate speed, momentum, spin, and dozens of other variables before each application of the brakes.

Unfortunately, its embedded processors can only perform one task at a time. To get around the problem, Associate Professor Stephen Edwards said, programmers slice tasks into many little pieces and have the processor hop between slices at such blindingly fast speeds, it presents the illusion of simultaneity.

This illusion—called concurrency—comes at a cost. Most programmers use the C language to code embedded processors. This involves lots of repetitive programming, and errors can creep in. “It’s like writing a phone book by hand. People could do it, but there would be lots of mistakes,” Edwards said.

Second, concurrent C programs are hard to test. C programs must be translated, or compiled, into the ones-and-zeros language of processors. Their sliced-up nature makes them hard to translate, model, and test. “The only way to tell if they will run fast enough to handle critical calculations is to test and retest programs until they appear to work,” he said.

Edwards has solutions for both issues. First, he has developed a language, Software-Hardware Integration Medium (SHIM), which simplifies programming concurrent events. SHIM reduces the errors that creep into repetitive programs. “We developed algorithms that automate all the bookkeeping necessary to manage simultaneous events,” he explained.

He also created a customized compiler that generates testable code. It takes C-like programs and translates the concurrent parts of the programs into a (very long) series of sequential commands. It then re-compiles them back into C.

“This lets you test your program in a model to check its speed and reliability and make improvements,” Edwards said. “This will lead to more reliable behavior and maybe fewer large recalls when embedded processors fail.”

“Embedded processors hide in the environment. As hardware plummeted in price, it became possible to put them everywhere. I ask students how many processors they own. They may count their computers or smart phones, but miss their coffee makers, air conditioners, and cars. We made 10 billion embedded processors in 2008.”

B.S., California Institute of Technology, 1992; M.S., California (Berkeley), 1994; Ph.D., 1997
Fruits and vegetables are perishable inventory because they spoil if grocers cannot sell them. The same is true of hotel rooms, rental cars, and airplane seats. Unless they are filled by a certain time, these services cannot produce revenue.

Corporations have become adept at selling perishable inventory by varying prices and running sales. These adjustments are called dynamic pricing, and Professor Guillermo Gallego is one of the field’s pioneers. He originally explored how customers value such attributes as a flight’s departure time, stopovers, seats, and luggage policies. His work is embedded in many of the models used to price perishable services.

Today, Gallego is working on “service engineering,” a concept similar to financial engineering. “It is similar to selling options on a stock,” he explained. “We take a basic service and create derivative services from it. This can be a win-win for buyers and sellers, and could dramatically change how certain services are sold.”

An example is a fulfillment option. Ordinarily, airline customers buy seats on specific flights. Gallego proposes that airlines offer a discount to customers willing to fly within a certain time period, say 9 am to 3 pm, and allow the airline to pick the flight. “The buyer gets a discount. The company buys flexibility, so they can accommodate business customers who often book late and must pay a premium price for a ticket,” he said.

Callable products are another possibility. Here, the seller discounts a service in return for the right to buy it back at a premium. A concert promoter, for example, might do that if a band becomes wildly popular during a tour. “In exchange for a discount, the provider can take advantage of a rise in prices,” Gallego said.

Gallego is also assessing options where consumers pay an up-front fee that gives them the right to buy a service, such as a hotel room, at a discount in the future. “Companies can afford to do this because not all consumers will exercise their options,” he said.

“There’s an art and science to engineering and pricing services, but it is always easier when they are win-win for buyers and sellers,” Gallego said. “I’m using service engineering concepts to help Hewlett-Packard redesign its warranties. We can offer annual warranties or month-to-month warranties that customers can drop at any time. Priced right, monthly warranties offer value to customers who replace products frequently and are profitable at the same time.”

B.S., California (San Diego), 1980; Ph.D., Cornell, 1988
Silicon, the second most abundant element in the Earth’s crust, is the key material of the modern information age. Microelectronic chips use bulk-silicon wafers to power computers, and silicon is used for increasingly important electronic applications, such as inexpensive solar cells, high-resolution flat-panel displays, radio-frequency identification tags, and threedimensional integrated chips. But manufacturers need high-quality crystalline silicon films in which atoms are nicely and periodically arranged. While it’s easy to obtain amorphous silicon films, they are not well-suited for making these electronic devices. Developing efficient ways to generate high-quality silicon films is a key to the proliferation of these micro- and macro-electronic applications. Professor James Im’s process for developing high-quality silicon film is playing a crucial role in developing the latest generation of flat-screens for a wide array of electronic devices.

Im has done extensive research that investigates how silicon, solid thin films and nano-scale structures behave when these materials are rapidly heated by laser irradiation, melted, and then subsequently solidify. While his studies look primarily at the scientific and fundamental issues involved, the findings also have led to various technical approaches for realizing high-quality silicon films on various technologically important substrate materials such as glass or plastics.

These laser-induced and melt-mediated crystallization processes, which convert initially amorphous or defective silicon films into low-defect-density silicon films, take place at temperatures above 1400 degrees C. According to Im, understanding how silicon melts and solidifies under these extreme conditions is critical for understanding how the atoms are subsequently packed and positioned.

“Knowing the fundamental details of how Si melts and solidifies makes it a rather straight-forward exercise for us to come up with efficient and effective ways to generate useful materials with periodically arranged atoms that make good electronic devices,” says Im.

The fundamental findings and technical approaches generated at Columbia are powering the evolution of the field of thin Si-film based electronics. One method, called Sequential Lateral Solidification (SLS), is used to manufacture high-resolution LCDs, and has recently emerged as the leading method for the next generation of flat-panel TVs, which use organic LEDs.

Top display makers, including LG Display, Sharp, and Samsung, have already licensed this technology. The innovation is also applicable to smart cards, RFIDs, image sensors, and three-dimensional integrated circuit devices.

In addition to laser-based approaches, Im is also investigating other beam-induced crystallization techniques that could provide unconventional, yet effective solutions for various electronic devices and applications.

B.S., Cornell, 1984; Ph.D., Massachusetts Institute of Technology, 1989
Some prices never sit still. Retailers discount clothing and technology products seasonally. Traders bid stocks up and down daily. Airline and hotel prices fluctuate by the hour. These are examples of dynamic pricing, where companies price goods based on cost, customer behavior, and competitive dynamics. Associate Professor Soulaymane Kachani’s research in the field has taken him in some interesting directions.

“We are applying traffic flow theory used in transportation networks to blood rheology to prevent blood clots,” he said. “Existing models are hard to calibrate for elderly patients because they require too much ultrasound data. Our models are simpler, and appear to better predict where clots will form. Our next step is to conduct clinical trials.”

One recent project assessed lifecycle pricing for different generations of technology products. He found that to maximize long-term profits, companies should not discount old technology too deeply.

“Those companies interact repeatedly with their customers,” Kachani explained. “Once they set a price, it affects the reference price. So if they start driving down the price of older goods, they cannot go back and ask for a dramatically higher price for their next-generation product. In fact, long-term, many tech companies are better off discontinuing old products than discounting to sell off inventory.”

Real estate, on the other hand, could benefit from more dynamic pricing.

“Imagine you’re developing condominiums,” Kachani said. “What price do you assign each unit?”

“You don’t want to sell out all the units with upper floors, good views, or two bedrooms first. If you do that, it means you did not put the right premium on the more desirable units. If the premiums are set correctly, all your different units should sell at roughly the same pace,” Kachani explained.

To find the right premiums, Kachani looks at both unit sales and also what units visitors view. He uses their actions as input for a computer model that modifies prices based on real market input. This gives developers a realistic way to set prices to maximize returns.

His work also extends to fashion. Kachani compared retailers who emphasize innovation and design with those who focus on pricing. The innovators, with short product runs and high turnover, had higher profits than retailers with larger product runs who relied on periodic discounts to clear the shelves.

Yet Kachani urged the innovators to consider dynamic pricing. “They would do even better if they managed their pricing strategy better,” he said.

B.S., École Centrale Paris, 1998; M.S., Massachusetts Institute of Technology, 1999; Ph.D., 2002
Software systems have a complex lifecycle, and Professor Gail Kaiser likes to work on all aspects of it. Her research ranges from creating systems that make recommendations to finding flaws in “non-testable” programs. “I like to find solutions in one domain and then generalize them,” she said.

For example, she is working on three “recommender” systems. One system monitors how biologists use tools for genomic analysis, and then gives novices recommendations based on the workflow of more experienced users. A second mines past experience to help programmers convert software to parallel code for multicore processors. The third helps computer science students solve certain errors in the code they write.

“We built all three systems independently, then noticed that they all used essentially the same architecture,” Kaiser said. “This lets us derive a general reference architecture that might be useful in building future recommender systems.”

Kaiser is also interested in testing so-called “non-testable” programs. These include machine learning, simulation, data mining, optimization, and scientific computing systems. “I come from a software engineering background, but work at the borders of my discipline and operating systems, databases, and security. I’m concerned with how to build systems – not just coding software, but how to design and test systems over their full lifecycle.”

“Ordinarily, you can look at a program’s input and see if the output is correct,” Kaiser said. “But what if you can’t? After all, non-testable programs are written to answer questions whose solutions are unknown. What if the answers are wrong in some cases, but not others? These programs could have all sorts of arbitrary errors, but how would we know?”

She has developed a number of approaches to test machine learning programs. One is a technique many math students will remember. Kaiser provides a problem, then changes the order of inputs to see if the program still generates the same answer.

“We found a lot of bugs in certain packages widely used in the machine learning community by using these approaches,” Kaiser said.

She has also developed methods to test for errors on computers deployed in the field. Field tests look at the widest possible range of software operating conditions. With so many variations, she can find subtle errors that elude even the most comprehensive lab testing programs.

“You can never get all the bugs out of them, but the more bugs you remove, the better.”

B.S., Massachusetts Institute of Technology, 1979; M.S., 1980; Ph.D., Carnegie Mellon, 1985
“The barbarians are no longer at the gates,” Associate Professor Angelos Keromytis said about computer security. “They are inside the doors and there are not enough guards to repel them.”

Most security systems are designed to keep bad guys out, and can do little once they are inside, Keromytis explained. “We start with the proposition that attackers will compromise your system, despite your best efforts to keep them out. The only solution is to make systems that are self-healing and self-protecting,” he said.

Keromytis’ approach is to teach computers to act like the best human experts, if they had all the time in the world to react to an attack. “We want the computer to recognize an attack, see what happens, and come up with a way to modify the system so that it blocks the attack,” he said.

Most attackers take advantage of the fact that nearly all computers on a network run the same software. If an attacker finds a vulnerability in one computer, it can attack all the computers. Keromytis turns this into an asset. His software monitors each system, noting when attacks fail or succeed and looking for unusual behavior.

When the alarm sounds, his security system isolates the infected computer. Then it analyzes recent events to find the trigger – an e-mail virus, a malicious download, a tainted document – that set it off. The system automatically attempts to write software code to fix the problem, testing different approaches until it finds one that works. It then rolls back the computer to a time before the attack and inserts the fix. The entire process takes only fractions of a second.

The newly inoculated computer also passes information about the threat around the network. Each computer then builds its own fix. This build-your-own approach prevents hackers from somehow attaching viruses to fake fixes.

“What we’re trying to do is build systems where the individual computers and servers collaborate to prevent attacks, fix attacks that succeed, and then send information to other parts of the network about the vulnerability so they can fix it too.”

Keromytis is currently testing the software and plans to scale up to larger systems soon. He is also looking at ways to find viruses that wait weeks or months until erupting.

The barbarians may have gotten through the gates, but in the future they will find the doors barred by a new generation of persistent guards.

B.S., Crete, Heraclion, 1996; M.S., Pennsylvania, 1997; Ph.D., 2001
While science has always been driven by data, the amount of data now available is growing exponentially, with data volumes doubling annually; that is, each year a community adds an amount of data equal to all that it previously possessed. Projects already underway are expected to generate even more data, and it will be generated at ever faster rates.

All disciplines are experiencing the flood of data – from the biological sciences and the physical sciences to the social sciences and the business world. Web-based data and government data continues to grow, but today’s computer systems can be overwhelmed by the immensity of the data flow.

Computer simulation has been used effectively to analyze the workings of the natural world and human-engineered systems in ways far less costly than experimentation. But now the massive data sets have overwhelmed those systems.

Professor David Keyes has developed scalable algorithms that allow computers to analyze a wide variety of operations – from the workings of the tokomak that contains the superhot reactions of up to 200 million degrees in magnetic fusion, to the problem of understanding the folding of a protein, which takes place in the span of a millisecond.

The ability to understand these massive data sets, and integrate with existing data, requires new strategies in what is called “data-enabled science.” Keyes now works on the frontiers of computing in data-enabled science, which allow computers to mine these huge data sets.

“The real excitement in science today is combining simulation and data-enabled science,” says Professor Keyes. “It’s a fresh frontier, where you are turning up the power of simulation, as you confront real data, in real time. It’s one of those grand challenges.”

In climate modeling, for example, Keyes (together with Columbia’s Haim Waisman in Civil Engineering) supports the work of climatologists analyzing the cracking of the polar ice shelf, where 77 percent of the world’s fresh water is locked up in land ice in Greenland and Antarctica.

Global warming has led to significant cracking of the ice, causing huge chunks to slide into the sea. Keyes is now working with the U.S. Department of Energy to describe previously ignored phenomena for their climate models, to generalize finite element models to handle cracks in the ice, and predict what sort of fractures are likely.

Keyes is a member of the NSF Advisory Committee for Cyberinfrastructure, and the NSF Advisory Committee for the Directorate of Mathematical and Physical Sciences.

B.S.E., Princeton, 1978; Ph.D., Harvard, 1984
According to Assistant Professor Martha Kim, the typical computer processor is like the family van. It is a good all-around machine, but it achieves its flexibility by sacrificing power and performance. It will get you there, but it’s not made for off-road adventure or hugging the curves at Le Mans.

Modern computers have the potential to act as vans, sport cars, and motorcycles—all on the same chip, Kim said. It is just a matter of getting under the hood and adding accelerators, small chips-within-a-chip designed to process certain types of data very efficiently.

Accelerators are possible because today’s chips are so large. Many contain several separate processors and upwards of 2 billion transistors. “We have transistors to burn,” she said. “If we could organize a few hundred thousand of these transistors into a specialized accelerator, we could handle certain types of data 100 times faster and with 100 times less power.”

Those gains are possible because everything needed to process information would reside in the accelerator. “Instead of reading and decoding software instructions specifying how to manipulate the data, the accelerator could start processing immediately without waiting for software instruction.”

“Today’s integrated circuits offer an embarrassment of transistors. The challenge is how to translate efficiently these raw resources into easy-to-use, high performance, low power processors. Spending some transistors on special purpose data processors, which store and manipulate structured data types, could simultaneously boost performance and conserve power.”

Accelerators are relatively simple to define in hardware but can be very difficult to use in software. “With accelerators, the programmer not only has to write the code, but has to coordinate what parts of the program should run on which accelerators, and then reassemble the results. Also, if the number and type of accelerators differs from chip to chip, programming becomes even more complex” Kim explained.

Kim’s goal is to create common interfaces and tool chains to protect the programmer from this complexity. “The programmer would write code normally. A compiler would track the libraries and data structures used by the application as well as the accelerators available on the chip. It would do the job of matching parts of the computation with the available accelerators,” she said.

In other words, Kim does not want to make her van into a Formula 1 racecar. She wants to keep all of the van’s flexibility and still take those turns at ridiculously high speeds.

B.A., Harvard, 2002; M.E., Università della Svizzera Italiana (Lugano, Switzerland), 2003; Ph.D., University of Washington, 2008
Many blame structured financial instruments, such as credit default swaps and collateralized debt obligations, for the 2008 recession. Yet similar products traded for decades without problems. Even now, billions of dollars in structured debt trade daily. Professor Steven Kou has made it his mission to make these products safer.

“As an engineer, I’m interested in linking economic theory to real-world pricing of structured financial products. Economists understand the structure of economic forces, and statisticians understand how one event triggers another, like the aftershocks of an earthquake. We’re trying to apply both to the details of financial products.”

Structured instruments reduce risk, Kou said. He points to instruments that pool corporate bonds. Instead of buying a bond from one company, investors can buy a diverse portfolio of bonds from companies in different industries.

Financial firms typically divide this basket of bonds into risk categories, or tranches. The top tranche has the lowest risk but the lowest returns. It loses money only if 30 percent of the bonds default. This is highly unlikely, and it trades like a highly rated bond. The bottom tranche has the highest return but loses money if only a smaller percentage of bonds default.

“There’s a value to this,” Kou explained. “Pension funds, for example, cannot invest in bonds rated less than AAA. Many strong companies have lower credit ratings. If their bonds are included in the top tranche, a pension fund can buy them without great risk and still receive a higher return.”

Many investors were lured by that combination of higher returns and lower risk. They believed that even if conditions in one industry forced a company to default, diversification would keep their investments safe.

In 2008, though, that assumption was upended. “The model we had been using was no good. During a severe crisis, we found that when one company defaults, others outside its industry are more likely to default,” Kou said.

Kou calls this “default clustering.” To understand how it affects risk and value, he builds models that draw on both economics and financial engineering.

Kou says the models will help set more realistic prices for structured financial instruments. Initial results are promising. Just before Lehman Brothers went bankrupt, conventional models set the cost of insuring the top tranche of corporate bonds at about $7,000. His model priced it at around $52,000.

“That’s more consistent with what happened in the market,” he said.

M.A., Columbia, 1992; Ph.D., 1995

Linking Domino Theories to Real-World Pricing

STEVEN S.G. KOU

Professor of Industrial Engineering and Operations Research
Automobile collisions account for tens of thousands of fatalities in the United States annually. While the most expensive automobiles have on-board collision avoidance systems, such technology is priced out of the market for most drivers. The cost has much to do with the technologies that are currently used to implement these systems. Current sensors rely on multiple integrated-circuit chips based on compound-semiconductor technologies, resulting in systems that are large, bulky, power-inefficient and expensive.

Silicon-based millimeter wave technology could make automobile collision avoidance systems as common as seatbelts in the cars of the future. Millimeter waves deliver good directionality, and offer a large amount of available bandwidth not currently being used, making them functionally comparable to fiber optics without the financial and logistical challenges. Silicon-based technologies offer the opportunity to integrate complex sensors onto a single chip, greatly reducing power, cost and size. This technology’s utility is wide ranging and includes collision warning systems, blind spot analysis, and pedestrian detection. It also is being explored for high-data-rate personal area networks for future “wireless homes,” non-invasive medical imaging, airborne chemical sensing, and concealed-weapon detection for security systems.

Integrating extremely high frequency electronic circuits and systems into silicon-based technologies is one of the grand challenges of electronics, and where Professor Harish Krishnaswamy is applying his research efforts. He pioneered silicon-based, nonlinear, multifunctional circuits and systems which, when coupled with millimeter wave technology, allow multiple simultaneous functions to be performed on a single, compact, power-efficient chip. A nonlinear, multifunctional phased-array transceiver chip won the prestigious Lewis Winner Award for Outstanding Paper at the 2007 IEEE International Solid-State Circuits Conference. Professor Krishnaswamy is also working on new Multiple-Input, Multiple-Output (MIMO) radar concepts that use multiple transmitting and receiving antennas to capture a more detailed and accurate image of the scene around the vehicle.

The Krishnaswamy Group at Columbia University analyzes, designs, and experimentally verifies novel integrated devices, circuits and systems for a variety of radio frequency and millimeter-wave applications. His research efforts blur the boundaries between circuits, electromagnetic device physics, and communication/signaling processing theory. Results include a variable-phase ring-oscillator based architecture for radio-frequency (RF) and millimeter-wave phased arrays, architectures and circuits for single-chip MIMO radar, timed arrays for ultra-wideband beamforming, and high-performance RF and millimeter wave building blocks for wireless transceivers.

B. Tech., Indian Institute of Technology (Madras), 2001; M.S., Southern California, 2003; Ph.D., 2009
During Professor Longman’s sabbatical in 1984 he initiated research in three new fields, becoming one of the very early contributors to each. With support of ex-doctoral student Dr. Robert Lindberg at the Naval Research Laboratory, an Egleston Medal recipient, he started research on robotics in space. The shuttle arm can handle a load of mass similar to the shuttle, and this creates a question: which end of the arm is the base and which is the load? Two of his early papers appeared as the first two papers in the first book on space robotics produced by the Carnegie Mellon Robotics Institute.

With German collaborators, he started research on time optimal control of robots, something that challenges numerical solution methods. One research focus was a press chain on the Mercedes production line near Stuttgart. The objective was to increase productivity by making the slowest robot get its job done faster. A series of publications progressed from idealized investigations to ones including detailed hardware constraints. Similar productivity problems appear in the production of semiconductor chips.

When a robot is commanded to follow a trajectory, it will repeatedly follow a somewhat different path. Robots often do the same operation hundreds of times a day, making the same errors each time. Professor Longman considered this a bit stupid – can’t we make a control system that learns from its experience to do what we ask? He started work on this at the University of Newcastle in Australia. Since then, this problem has developed into the fields of iterative learning control (ILC) and repetitive control (RC).

Professor Longman has produced approximately 250 publications in this area, and is known for advancing the theory in a way that produces improved real world performance. Experiments on a robot at NASA improved tracking accuracy by a factor of 1000 in just 12 iterations for learning. The methods can apply to a very large number of feedback control systems, creating high precision motion by improved algorithms instead of higher precision hardware.

At Seagate Technologies, experiments reduced the repeatable error in computer disk drives by 98 percent. Similar experiments improved paper handling in copy machines at Xerox. Experiments also demonstrated improved beam focus at the 8 GeV (one thousand million electron volts) accelerator at Jefferson National Accelerator Facility. He is currently working on similar experiments at the Naval Postgraduate School on jitter control in laser optics on spacecraft.

ILC and RC aim for high precision motion and optimal control aims for fast motion. Professor Longman is working to develop a marriage between these research areas to simultaneously get the benefits of both – aiming for higher quality products created with improved productivity.

B.A., California (Riverside), 1965; M.S., California (San Diego), 1967; M.A., 1969; Ph.D., 1969
Cloud computing—delivering software and services from a central computer to
desktop terminals—is arguably the hottest topic in computing today. The rea-
sions are economic. PC hardware prices continue to fall, but maintenance costs
continue to rise.

“If you’re a large corporation with 50,000 or 100,000 desktops, you’re fixing bro-
ken hardware, guarding against viruses, and patching and upgrading software for each
one of them. The costs are astronomical,” explained Associate Professor Jason Nieh.

If the PC-on-a-desktop paradigm is broken, what will replace it? “Most analysts
believe we are moving to cloud computing, where corporate computers run only in se-
cure data centers where they are protected, secure, and easier to manage and service. If a
desktop fails, it doesn’t matter because all the memory and files actually reside in the data
center,” said Nieh.

But cloud computing has a weakness: speed. Centralized applications run slower
than the same program on a local PC. This is especially true for programs with graphical
displays.

“A modest display has 1024 x 768 pixels, and each pixel has 32 bits of data. Dis-
plays update 30 to 60 times per second or more, so you’re potentially sending a gigabyte
or more of data per second to each PC on the network, and that can slow response
times,” Nieh said.

Computer scientists have tried to compress data to reduce the load. This helps,
but it requires additional computing power and fails to handle gracefully today’s complex
graphical interfaces.

Nieh uses intelligent software to reduce data flows and response times from the
cloud. In Nieh’s scheme, the application draws the screen on a virtual display. Then his
program analyzes what is on the display, and sends commands to the desktop terminal,
instructing it on how to redraw the screen. Many of the most common commands are
embedded in the graphics card’s hardware, so they operate very fast. The system updates
the terminal by sending only those portions of the display that change, enabling very fast
response times.

The big payoff comes when connecting to the Internet. Data centers almost
always have the fastest Internet connections. “They update Web pages much faster than
local desktops, laptops, or smart phones,” Nieh said. “If a carrier uses this technique, you
don’t have to settle for the limited functionality of smart phones that run some software
but not others,” Nieh explained.

“You get improved functionality and improved performance that makes you feel
like you’re right there, and you get it on your smart phones, desktops, and laptops.”

B.S., Massachusetts Institute of Technology, 1989; M.S., Stanford, 1990; Ph.D., 1999

Delivering Desktop Computing from the Cloud

JASON NIEH

Associate Professor of Computer Science
For decades, computer processors were typically organized like marching bands: a conductor kept time and band members stepped to the beat. In processors, a clock’s pulse determined when all computations and data movement occurred. Today, that paradigm is breaking down, Professor Steven Nowick explained.

Modern processors consist of a handful of smaller processors, or cores. “When you have four separate cores, it is difficult for one clock to keep them in lockstep,” Nowick said. The problem will only worsen when future processors have dozens of cores.

Today’s transistors also pose problems. As they shrink to a few tens of nanometers, they become much more variable. “Their speeds vary depending on temperature, voltage, and how they are manufactured. Their unpredictability is a major design challenge,” Nowick said.

Nowick and colleagues at other institutions have been pursuing an alternative approach: eliminate the clock and let digital components operate at their own speeds. “Let them communicate as conditions require, and make their own decisions with their neighbors about when they need new data and when they will output results,” he stated.

“Most digital systems have clocks running at billions of cycles per second. Everything operates in lockstep with that clock. As circuits get larger and more complex, imposing fixed timing on billions of transistors and millions of components is a huge design effort. We think we can solve these problems with asynchronous, or clockless, circuits.”

It sounds chaotic, but the Internet works the same way, Nowick said. “People around the world add, update, and remove Web pages individually, without any centralized control mechanism.”

In addition to solving timing issues, asynchronous digital systems could provide other advantages. In synchronous chips, even idle components are activated every clock cycle, like band members marching in place. In contrast, the on-demand components in asynchronous systems respond only when necessary. This conserves energy and can prolong battery life in laptops, smartphones, and other portable devices.

Asynchronous processors are potentially easier to design, since new circuits do not have to be synchronized with the entire chip. "It’s a Lego-like system, which can be snapped together," Nowick said. Hurdles remain. Engineers need new software tools to design asynchronous circuits, and face subtle issues in designing these circuits correctly.

Nowick is currently working on both challenges, including projects to design a flexible asynchronous interconnection network for future desktop parallel computers, and ultra-low energy signal processors for hearing aids and medical implants.

When disaster strikes, the interdependent complexity of the environment (utilities, transportation, and communication infrastructures, homes, and office buildings) can result in a cascading effect that quickly exacerbates the crisis. Large-scale disasters, such as Hurricane Katrina and the earthquake in Haiti, have graphically demonstrated the need for reliable initial disaster preparedness, response, and recovery. In such cases, the immediate availability of critical real-time data is crucial to saving lives.

Feniosky Peña-Mora, dean of The Fu Foundation School of Engineering and Applied Science at Columbia, has developed a new disaster response framework—Collaborative Preparedness, Response and Recovery (CP2R)—that makes a significant difference in the outcome of such disasters. As part of this framework, he and his research team have created a mobile workstation using an all-terrain, heavy-duty Segway personal transporter outfitted with a payload that can include a Tablet PC, infrared and thermal still and video cameras, Global Positioning System receivers, and other advanced data collection technology. These instruments can collect, archive, analyze, and report large quantities of data to provide better situation awareness of an emerging disaster response scenario, and automatically generate digital models that can be used for disaster response.

By deploying these modified chariots manned by civil engineers, real-time data from first responders can be transmitted to coordination centers by wireless voice and data communication infrastructures. “This new cohort of first responders will provide accurate, real-time information to support technically sound decision-making processes during both the initial disaster response and the recovery phases,” says Peña-Mora.

“With a legion of mobile workstation chariots, we will be able to mitigate the dynamics of the disaster by improving the dynamics of the disaster response.”

Recent testing of the mobile chariot has shown the potential for its success in the field. Despite additional weight from mounted instruments, the unit retained its stability on uneven surfaces and in differing weather conditions. Using digital images collected at the disaster site, decision-makers in coordination centers can evaluate infrastructure stability, study how the first responders are reacting to changing situations, and collect data for future analysis.

Dean Peña-Mora holds appointments as professor of civil engineering, computer science, and earth and environmental engineering, is the author or co-author of more than 150 scholarly publications, and holds five patents, one provisional patent, and one technology disclosure.

All computer software has one thing in common: it runs on computer hardware. But what if you could not trust the hardware to securely run software? That's the question posed by Assistant Professor Simha Sethumadhavan. "All software runs on hardware. If the hardware is hacked, then it can subvert all software and software security countermeasures. Since hardware is the root of trust, attacks on hardware are potentially very dangerous," Sethumadhavan said.

Until recently, computer scientists never suspected that someone could tamper with hardware. Yet investigators have found unusual additions in military chips. One way to prevent hardware hijackings is by passing tokens every time data moves within hardware. Sethumadhavan likens this to sending a thank you card after a gift. "Let's say Charlie wants to contribute $100 to Alice's charity, but has to send it to Bob first," he said. "Bob takes $10 for himself and pays the rest to Alice. One way to find out if there is a problem is for Alice to write Charlie a thank-you note for the $90 donation. When Charlie sees the discrepancy, he asks accountants to trace the missing money."

Sethumadhavan proposes creating similar triangle-like structures within a computer processor. "They would monitor any irregularities. We want to create a chain of monitored data and sound the alarms if any of the links break. These lightweight monitoring additions incur very little processing overhead," he said.

"We are taking a clean slate, ground-up approach to designing secure systems. As a foundational step, we have designed methods to protect processors, the core of all computing infrastructure. Once processors are secure, we can securely build out support for protecting other hardware and software."

Sethumadhavan is also working on other techniques for securing processors. "All hardware back doors have triggers and payloads. The triggers are usually time or data input values that activate the payload," he said.

"We are working on ways to silence the trigger. For example, we might be able to reset the processor's counter so it never reaches the threshold value needed to trigger an event. Or we could use lightweight encryption to obscure data values. Only when we fully trust our processor can we fully trust other security procedures," Sethumadhavan said.

Sethumadhavan is leading a project on rethinking security, making it a priority instead of an afterthought, with three other Columbia Engineering professors and a team from Princeton University. The project, titled "SPARCHS: Symbiotic, Polymorphic, Autonomic, Resilient, Clean-slate, Host Security," just picked up a federal grant for more than $6 million.

B.S.E., University of Madras, 2000; M.S., University of Texas, 2005; Ph.D., 2007
Finding a Fairer Way to Admit Students

Jay Sethuraman

Associate Professor of Industrial Engineering and Operations Research

Associate Professor Jay Sethuraman began his career by matching sets of jobs with machines to improve factory performance. More recently, though, he has used operations research to find the fairest way to admit students to top public high schools.

To Sethuraman, the two problems are similar, with one major exception: “In a factory, the machines don’t care what job they do. But schools do care about which students they admit,” he said.

New York City’s selective schools choose students based on admissions test scores. Students who do not get into their top choice can appeal. In fact, schools may set aside a certain number of seats for appeals, in addition to those seats lost when students they admitted leave for another school.

Resolving appeals fairly and efficiently discourages students from gaming the system, Sethuraman said. They may be willing to go to several schools, but list only one school if that increases their chance of placement. Or they may list schools that are unlikely to admit them if it improves their odds of getting into the school they want.

“A better system would give students an incentive to list their true preferences without penalizing them for doing it,” Sethuraman said. “We want to maximize the number of students who get into their top choice, but treat all students in a fair, systematic way.”

Under Sethuraman’s approach, each student starts with a seat in a school that he or she wants to trade. Rather than trade individual seats, students exchange their seat for a fraction of a seat in the schools they want to attend. Those fractions, which add up to a full seat, are computed based on seat availability plus the desirability of the student’s existing seat.

“At the end of this procedure, a student may have one-half a seat in school A, one-third a seat in school B, and one-sixth a seat in school C. This determines their probability of getting a seat in the lottery,” Sethuraman said. Students who do not complete a trade move onto the next round of lotteries, where their odds are reset to account for the remaining available seats.

“Listing all the schools you are willing to attend increases your chances of staying in the game longer and getting into a school you really want. We give students an incentive to list all their acceptable schools without trying to game the system,” Sethuraman said.

B.E., Birla Institute of Technology and Science (India), 1991; M.S., Indian Institute of Science. 1994; Ph.D., Massachusetts Institute of Technology, 1999
When people refer to the World Wide Web as an information superhighway, they rarely consider traffic jams. Yet congestion slows the movement of information around the Web, and appears naturally in systems as diverse as highways and hospitals.

Professor Karl Sigman uses probability tools to build and analyze mathematical models of congestion, also known as queueing. A simple example is an ATM machine, where people arrive randomly and sometimes find themselves waiting in line to use the machine. Successfully analyzing queueing models can help optimally route requests to a set of Web servers, staff a call center, process jobs in a manufacturing plant, and schedule surgeries in a hospital.

The mathematics of probability gives Sigman many insights into a model’s evolution. Still, many models remain breathtakingly complex due to the inherent randomness involved in the real world.

“Randomness, such as when the next request arrives or when something breaks, affects all these systems. The further you look into the future, the more random it can become. It’s like stock prices. Tomorrow’s price is likely to be similar to today’s, but the price next week is less certain,” Sigman said.

“In 2000, our team used Internet polling data from all 50 states to simulate the Bush-Gore election. We ran one million probabilistic simulations. Bush was ahead in the national polls, but our simulations showed that Gore was likely to win the Electoral College. The only thing we got wrong was Florida.”

“I’m interested in the relationship between what system users see and what the system actually does,” he explained. “A user might click a link on a website. How long he or she waits to see the page is a measure of congestion from the user’s perspective.”

A system observer’s viewpoint is different. “He or she looks at the Web server over time and asks, ‘How many users are trying to access a given page?’ It does not look at the experience of any given user,” Sigman said.

“This is also a measure of congestion and system performance, but from different perspectives,” Sigman added. Yet the two views are interrelated. In fact, the solution to a problem from one perspective can sometimes be transformed into the desired solution from the other perspective.

Sigman has spent years teasing out those connections. “Sometimes the model looks very complicated from the perspective of a user, but it proves easier to solve from the perspective of an observer,” he said.

B.A., California (Santa Cruz), 1980; M.A., California (Berkeley), 1983; M.S., 1984; Ph.D., 1986
Aging infrastructure is a major problem around the world and monitoring the health of structures, from bridges to dams to buildings, is critical to our modern society. Andrew Smyth, professor in the Department of Civil Engineering and Engineering Mechanics, specializes in structural health monitoring, using the dynamic signature of a structure to determine its condition. This can include assessing a structure’s day-to-day performance, locating and quantifying potential areas of damage, or calibrating a model that can be stressed in a computer simulation for a heretofore-unseen loading event.

One of Professor Smyth’s recent projects has focused on monitoring vibrations on New York City’s Manhattan Bridge. To assess the bridge’s performance subsequent to a major retrofitting and strengthening program, and to calibrate a mathematical model of the bridge to predict its performance in the event of a potential seismic event, Smyth and his team placed a variety of different sensors that detected dynamic motions on the bridge over a two-month period. With the recorded data and their newly developed data fusion algorithms—a new technique that combines data from multiple sources—the team was able to identify the dynamic characteristics of Manhattan Bridge.

Smyth has also pioneered the use of differential GPS technology in conjunction with the data fusion technique to obtain highly accurate measures of low-frequency bridge deformations. He continues to develop data fusion algorithms for other civil and mechanical systems that combine information from a network of different kinds of sensors used to measure the dynamic response of a system. He says that, by taking advantage of the various levels of data redundancy, one can get high-fidelity virtual-sensing information that plays to the respective strengths of different types of sensors.

“Basically our work allows us to better understand the condition and performance of the built environment,” says Professor Smyth. “This really is our society’s most valuable physical asset and the backbone of our way of life. Structural health monitoring allows us to better allocate our resources to maintain and improve our infrastructure, and keep us safe.”

In 2008, Smyth was awarded the prestigious Walter L. Huber Civil Engineering Research Prize of the American Society of Civil Engineers. The award recognizes notable achievements by younger faculty members in research related to civil engineering. Professor Smyth was recognized “for fundamental contributions in the highly efficient identification and modeling of nonlinear deteriorating structural dynamics.” The selection committee commented that his research is characterized by “thoroughness, novelty, relevance, and intelligent breakthroughs.”

B.A./B.Sc., Brown, 1992; M.S., Rice, 1994; Ph.D., University of Southern California, 1998
If your credit card company ever called to confirm a purchase, you have entered Professor Salvatore Stolfo’s world. Stolfo specializes in detecting anomalies, events that stray too far from expected patterns. In addition to fraud, anomaly detection can be used to monitor engineered systems, sensor networks, ecosystems, and computer security.

Stolfo entered the field after inventing an algorithm that let marketers merge lists of consumers and purge bad records. “I realized I was aiding and abetting people who pierced personal privacy. It was an ethical dilemma,” he recalled.

His interest in privacy led to cybersecurity and eventually to the study of insider attacks. “Most security breaches are the fault of the humans. Someone didn’t implement something, or stole an identity, or had a grudge against an organization,” Stolfo said.

This differs from most security research, which aims to keep out hackers. University researchers are more ambitious, developing inherently secure programming languages and self-repairing systems. “These are important aspects of security, but they don’t matter if your adversary is already inside,” Stolfo said.

“There are many different types of insiders, and they all do things in different ways. We think of it as a chess game. What if insiders can control system access? If they can blind the system to their actions, they can get away with anything. We want to stop them.”

The most common type of insider threats is unintentional users. They may disable security measures to do their job more easily, or inadvertently push two buttons and erase a day’s work. “These are the most prevalent and least dangerous insiders,” Stolfo said.

Masqueraders include credit card thieves with stolen credentials. “The credentials make them insiders,” Stolfo said. He works with banks to model consumer transactions. “We’re always looking for ways to use more data to find problems sooner,” he said.

Maliciously intentful insiders use their own credentials to copy secret government or corporate documents, steal money, and even sabotage the system. Highly privileged insiders have a similar agenda, but they are the ones responsible for detecting other intruders.

To foil these intruders, Stolfo looks at how their behaviors vary from company norms. By plotting how users interact with software and documents, he hopes to find patterns that suggest malicious intent. He has also developed decoys to ensnare bad guys.

“Ultimately, we want to define metrics for what it means to be secure. Then we can start to build a science of security,” he said.

B.S., Brooklyn, 1974; M.S., New York University, 1976; Ph.D., 1979
While the transistor revolutionized the field of electronics and paved the way for personal computers, it made way for several perpetual challenges: deliver more power in smaller sizes, enable real-time interaction with the real world, and constantly adapt to technological change. Solutions to those challenges can make possible, for example, biomedical ingestible pills, containing chips that aid, or give information about, the body; sensor networks that provide information about the environment or physical infrastructure; or wireless communication technology that uses less battery power but provides more range.

One of the challenges in making this new era a reality lies in advancing the development of single silicon chips that perform both analog and digital signal processing. Analog and digital signal domains have significant technical differences, yet new technology demands more and more complex mixed-signal design. The development pace is relentless, driven by demands for increased performance. New techniques need to be invented, and fundamental limitations must be better understood, to make such analog/digital circuits with improved performance possible.

Dr. Yannis P. Tsividis, the Charles Batchelor Professor of Electrical Engineering, has been an important contributor to the field of silicon chips that mix analog and digital circuits. He and his students have done extensive research in this field at the device, circuit, system, and computer simulation level.

In 1976, Tsividis designed and built the first fully integrated MOS operational amplifier and demonstrated its use in a coder-decoder for digital telephony. These results were widely adopted by the industry in the first massively produced mixed-signal MOS integrated circuits, which incorporate both analog and digital functions on the same silicon chip.

Tsividis and his students have since been responsible for several important contributions, ranging from precision device modeling and novel circuit building blocks to new techniques for analog and mixed-signal processing, self-correcting chips, switched-capacitor network theory, RF integrated circuits, mixed analog-digital Very Large Scale Integrated (VLSI) computation and the creation of computer simulation programs. This work has resulted in several patents in several countries, and has been incorporated by the industry into products we use every day.

Dr. Tsividis is a Fellow of the Institute of Electrical and Electronic Engineering (IEEE).

B.E., Minnesota, 1972; M.S., California (Berkeley), 1973; Ph.D., 1976
Anyone who has driven on highways understands that random events affect congestion. Even in relatively light traffic, with no accidents or obstructions, cars will suddenly bunch up, slow, and then speed up again.

Professor Ward Whitt studies the enigma at the heart of this process. His discipline, queueing theory, examines how random fluctuations in flow, waiting, and processing cause congestion in complex systems.

Examples are everywhere. "We all spend too much time waiting on lines, from physical lines in a supermarket or bank to invisible lines on hold for a call center or waiting for a Web page to load," Whitt said. Queues are equally present in the waiting times of a computer processor or the movement of parts through a factory.

One major goal of queueing theory is to reduce waiting. Understanding congestion helps engineers specify the right number of telephone switches, Internet servers, and even call center personnel.

Despite their wide use, queueing models have a significant weakness. "The standard queueing models assume random flow, but the rate of that random flow is assumed constant. In reality, the arrivals to a system occur randomly, but the rate of that random flow is not constant," Whitt said.

Whitt tries to capture that systematic variation in the flow rate together with the uncertainty about that flow rate. He builds and analyzes models that reflect both these features of everyday queueing phenomena. "This produces high fidelity descriptions of congestion that go far beyond standard textbook queueing models," he said.

Whitt is also applying these insights to complex networks. "Queues do not appear in isolation, but appear in networked systems with multiple flow paths and queues," he said.

One way to tackle complex, networked systems is to see how they would behave as they scale up. "Sometimes," Whitt said, "a larger model tells a clearer story. Toss a coin 20 times and you expect to average 10 heads and 10 tails, but you may see from seven to 13 heads. But toss the coin one million times, you are likely to get closer to a 50-50 split."

He has developed mathematical techniques that show how congested systems behave at larger scales. He then compares the model with computer simulations of the system or data from that system.

"When you do this, you can end up with a fairly simple story that tells you a lot about your system," Whitt said.

A.B., Dartmouth, 1964; Ph.D., Cornell, 1969

Unraveling the Mysteries of Congestion
WARD WHITT
Wai T. Chang Professor of Industrial Engineering and Operations Research
Even the best written software contains errors. Assistant Professor Junfeng Yang wants to unmask and correct those often subtle defects. The software bugs are costly. In 2002, the National Institute of Science and Technology put their cost at $60 billion annually. Bugs do more than crash computers. They contributed to the northeast power blackout in 2003, and delivered lethal doses of radiation to hospital patients.

"My research involves finding ways to make software more reliable," Yang said. In graduate school, he developed an automated method to detect storage system errors. "Past tests were like throwing darts and hoping to hit a problem area. We developed systematic ways to test all possible storage states," he said.

After joining Microsoft, he extended his work to distributed storage systems on large networks. "People knew they were losing data, but not why. Our tool helped them find those bugs," Yang said. His work led to numerous patches for Microsoft's production systems and the Linux operating system.

Now Yang is focusing on the reliability of multithreaded programs. Unlike programs that run all their instructions sequentially, multithreaded programs consist of segments, or threads, that run concurrently. Multithreaded programs are significantly faster than sequential code.

They are also more difficult to write, test, and debug. "This is because they are not deterministic," he explained. In other words, a multithreaded program may behave somewhat differently each time it runs. "It may act correctly or buggy, depending on such variables as processor speed, operating system scheduling, and what data arrives when during operations," Yang said.

Lack of determinism makes it difficult to reproduce errors, much less fix them. Yang's research makes multithreaded programs execute deterministically, so programmers can isolate problems.

Explaining his approach, Yang likens threads to cars driving down a four-lane highway. "The cars drive in parallel lanes. During nondeterministic execution, they can change lanes whenever they want. When they do, sometimes they collide and cause the program to crash."

"To make threads execute deterministically, we've place barriers between the lanes. We only allow threads to change lanes at fixed locations, following a fixed order. This prevents random car collisions," he said.

Yang records this path and makes every subsequent group of cars follow it. "Because we know the path causes no collisions, there should be no collisions when another group of cars use it," he said. By attacking multithreading, Yang hopes to weave more reliable software.

B.S., Tsinghua University (Beijing), 2000; M.S., Stanford, 2002; Ph.D., 2008
What do hospitals, airlines, supply chains, and the Internet all have in common? According to Professor David Yao, they are all complex networks that must bring together multiple services and assets to accomplish any task. They must also share these same resources among different classes of customers, who pay different amounts for service.

Organizations want to manage their resources efficiently to maximize profits. But if they are too efficient – Yao likens it to filling a highway with cars so traffic slows to a crawl – they sacrifice quality of service. Balancing efficiency and service across complex networked resources is an exercise in extreme juggling. Yao wants to help by giving organizations the tools to do it in real time.

He points to airlines as an example. They must divide a limited number of seats among first, business, and several types of economy classes. Each class sells for a different price.

Airlines maximize revenue when they fill every seat. They can do this by discounting and by overbooking flights, since they know there will always be some no-shows. They also reserve some tickets to sell at higher last-minute prices.

That leads to problems. “The price they pay for overbooking is that they may have to ask people to get off the airplane. They also don’t want to hold too many last-minute tickets, or they will have unfulfilled seats,” Yao said.

Airlines estimate how many seats to sell and reserve by looking at past data. “That does not capture the real time dynamics of the network,” Yao stated. “On a particular day, a plane might be delayed and those passengers will need new connecting flights. Now their planes must carry their own customers plus passengers from the delayed flight.”

Yao’s models capture that type of real time information and use it to optimize the entire system rather than a specific resource, like a single flight, a bank of servers, or a hospital bed. On airlines, his models assign all seats a shadow price, the revenue they could potentially earn if they sold a reserved ticket, and compare it with the probability of delays and other events as they evolve. It shows them the most profitable way to reroute passengers and flights.

“We look at the probability of events, but also at how we can hedge our bets if that probability is wrong. We want to create models that are predictive but robust, so if you’re off, you won’t walk away from money on table,” Yao said.

Fudan University, M.A.Sc., Toronto, 1981; Ph.D., 1983

Professor of Industrial Engineering and Operations Research

Optimizing Networked Resources

DAVID YAO

Professor of Industrial Engineering and Operations Research
Dell, Yahoo, Google, and Facebook were founded by college students, Professor Yechiam Yemini tells his Principles of Innovation and Entrepreneurship class. He wants to teach students how to create innovative technologies and transform them into successful startups.

Yemini has combined academia with serial entrepreneurship. His first company, Comverse Technology, co-founded in 1984, revolutionized voice messaging technologies. Ten years later, System Management Arts created the first products to diagnose network failures automatically.

Startups, he explained, are another way of disseminating basic knowledge. “High-tech startups,” he explained, “distill the value in raw, basic technologies by creating innovative products and introducing them to the market,” he said.

Yemini’s course rests on three legs. The first is understanding how to identify opportunities. The most fertile areas are those where new ideas disrupt established ways of doing things, such as integrated circuits, the Internet, and wireless networks.

“Today, the biggest transition is from cellular phones to mobile computing. Now your phone is a tool to go shopping, access content, play video and read books. It’s a wonderful opportunity to launch companies that exploit this,” Yemini said.

The course’s second leg involves startup mechanics. “We look at the engines of value creation,” Yemini said. “Different engines make products, exchange information with the market, and manipulate the flow of financial resources. We look at how to design these engines to optimize the value they create while minimizing risks and errors.”

Yemini’s third leg is product development. “Many companies fail because they spend all their time creating a product and then look for a market,” he said. “They didn’t manage the risk that customers wouldn’t like their implementation, or that market needs might change.” He advises students to begin talking with customers from day one, and to keep improving products incrementally until they are happy.

Currently, Yemini is focusing on managing mobile services. “Mobility presents a disruptive change in delivering network services. It presents research opportunities to create new technologies, which may one day lead to new startups,” he said.

“A startup company is a bunch of engines that express the value of a technology. Think of it as a mechanism, a black box. There are ways to build a better box, ways to engineer it to better distill the value of the underlying technology. My course on innovation and entrepreneurship tries to teach how to engineer a technology company, much as one engineers other innovative constructs.”

B.Sc., Hebrew (Jerusalem), 1972; M.Sc., 1974; Ph.D., California (Los Angeles), 1978

Turning Students into Entrepreneurs

YECHIAM YEMINI
Professor of Computer Science
Computer chips are the building blocks that allow billions of transistors to fit in a small area. These chips have enhanced everyday life, and enable the design of electronics of increasing functionality and lower cost, making most modern day technology possible. But as transistors continue to become smaller and faster, new challenges for circuit designers constantly arise. The research field of Very Large Scale Integration (VLSI) addresses these challenges.

One of those challenges is transistor current leakage, which is becoming a bigger problem as transistors in computer chips continue to shrink, leading to problems with power and reliability. While current leakage and power dissipation in each transistor remain quite small, they can add up to a significant amount over billions of transistors, potentially limiting function and performance. Solving this problem could have a big impact on industry, and the feasibility of critical future applications of electronics.

Prof. Charles Zukowski, past chairman and current vice chairman of the Department of Electrical Engineering, has worked in the area of VLSI throughout his career and has contributed to the progress of integrated circuit technology in a number of areas. His chief focus now is twofold: circuit techniques such as monotonic logic to reduce the impact of current leakage in future integrated circuit technologies; and special-purpose hardware prototypes for the simulation of gene regulatory networks. Through this work, his intention is to further the capability of integrated circuit technology and to explore new applications.

His research has covered both circuit design and circuit analysis, results of which include a patented circuit technique for generating high data-rate serial data from a number of lower data-rate channels, and an approach for mixing digital and large-signal analog computation for simulation. He derived a number of results for bounding the behavior of digital integrated circuits that were compiled into a research monograph, and based on this work, he received an NSF Presidential Young Investigator Award. He later developed a technique for measuring the convergence of waveform relaxation algorithms for simulating digital circuits. He also proposed a technique for significantly reducing the power consumption in certain content-addressable memories and investigated the use of various memories and circuit techniques in internet routing hardware. Throughout, he has consulted for industry in the field of Complementary Metal Oxide Semiconductor Integrated Circuit (CMOS IC) design.

B.S., Massachusetts Institute of Technology, 1982; M.S., 1982; Ph.D., 1985

Plugging the Leak in Circuit Efficiency

CHARLES ZUKOWSKI
Professor of Electrical Engineering