UNIVERSITY OF CALIFORNIA, SAN DIEGO

UCSDCSE
Computer Science and Engineering

LETTER FROM THE CHAIR

In recent years the National Academy of Engineering (NAE) conducted an intense study of grand challenges and opportunities in engineering facing those born at the dawn of this new century. With input from leading thinkers, practitioners and even online public commentary, the NAE converged on four themes – sustainability, health, security and living – and 14 grand challenges. Six of these challenges are related to topics directly addressed by computer science and engineering as a discipline: health informatics, secure cyberspace, enhance virtual reality, reverse-engineer the brain, personalized learning, and tools of scientific discovery. So it should come as no surprise that the problems we tackle as computer scientists and engineers are increasingly inspired by big issues facing society: better education, better health care, better democracy, and better infrastructure for society's lifelines in energy, water and transportation.

It is truly an exciting time to be a (computer) engineer or a (computer) scientist! Enabled by data, our research is at the center of new sensing, learning and other paradigms as we enable new tools of scientific discovery through the lens of computing. The growing importance of our endeavor – from maintaining critical infrastructure to the secrets of the human genome – makes us humble. We must try our level best to understand how we can enhance education and learning outcomes in areas so critical to society and the economy. Our research covers a lot of ground, including theoretical computer science, computer architecture and networking, bioinformatics, databases, graphics and vision, machine learning, programming languages, software engineering, and much more. We also explore systems at multiple scales and dimensions: from the nano-scale to the macro-economic, and from embedded systems to computational biology. To do all this well requires a critical mass of talent, drive and place.

It is my intense pleasure to introduce you to that place: UC San Diego's Computer Science and Engineering department, one of the largest for computer science in the nation – with 550 incoming freshmen in Fall 2012! It is also the fastest-rising computer science department of the past decade. Our youth, compelling raw talent and tremendous chutzpah make us special: we attract the largest population of undergraduates, graduate students and research staff across all of our peers. Our size is exceeded only by the tremendous quality of our students, research and teaching staff. We take pleasure in producing students who are inspired and engaged, who can push the state of the art in knowledge and the practice of computing in institutions ranging from leading companies to the nation's top research universities. Of all things, I am most proud of the ecosystem of academic centers and research institutes that surround our students and faculty: the Center for Networked Systems (CNS), Center for Wireless Communications (CWC), San Diego Supercomputer Center (SDSC), and the California Institute for Telecommunications and Information Technology (Calit2). Together they provide a unique and compelling environment in which our researchers can work together and conduct research on projects that scale and provide ample opportunities for our talent to imagine, to discover and to build the future for computing and society.

We welcome this opportunity to tell you what we are all about. And if you haven't already, please plan to visit us in person to see for yourself the exciting confluence of bright minds, a beautiful climate, wonderful teaching and groundbreaking research!

Rajesh Gupta
Professor and Chair
Department of Computer Science and Engineering
Jacobs School of Engineering
University of California, San Diego

CONTENTS

Research Groups

CSE by the Numbers

56 Faculty Members • **8** Research Scientists • **12** Postdoctoral Researchers • **23** Staff
1,650 Undergraduates • **302** Graduate Students (**111** M.S., **191** Ph.D.)
#14 Worldwide ranking of UCSD computer science, according to 2012 Academic Ranking of World Universities
$25,679,949 Annual Research Expenditures (FY10-11)
5,079 Degrees awarded since CSE became a separate department in 1987

** Faculty and enrollment figures are for 2012-'13 academic year*

About the Cover

On the Cover: "Spectral Spiderweb" won First Prize in the 2011 competition among teams of students in CSE 168, the Rendering Algorithms course taught by CSE Prof. Henrik Wann Jensen. Undergraduates Bridgette Wiley, Jason Greco and Marlena Fecho tackled the spectral dispersion of light and used techniques including procedural generation (of the spiderweb), constructive geometry, Fresnel refraction and reflection, blooming, anti-aliasing, local tone mapping, and area lighting. For more information, visit http://bit.ly/XAgX3A.

UCSDCSE
Computer Science and Engineering

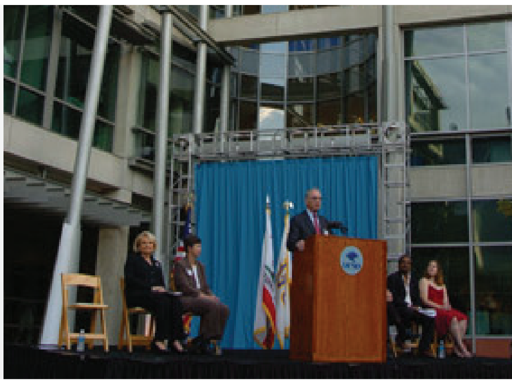
THROUGH THE YEARS

2013

2013
CSE establishes Moxie Center entrepreneurial incubator for undergrads with gift from Irwin Zahn

2010
First offering of Master of Advanced Study, with course in Wireless Embedded Systems

2010



2005
Dedication of 5-story, $41 million, 148,000-square-foot CSE Building

2005

cns
center for networked systems

JUL 2004
Center for Networked Systems founded under CSE Prof. Andrew Chien

2004

MAR 2004
Prof. Henrik Wann Jensen wins technical Academy Award for simulating subsurface scattering of light in translucent materials, e.g., to simulate skin. (Pictured immediately behind actress Jennifer Garner)

2000

2000
California Institute for Telecommunications and Information Technology (Calit2) founded under CSE Prof. Larry Smarr

1997

1997
School of Engineering named in honor of Irwin and Joan Jacobs

1987
Department of Computer Science and Engineering splits from Electrical Engineering with 18 faculty, 200 students

1987

SEP 1986
B.S. in Computer Science degree accredited by the Computing Sciences Accreditation Board

MAR 1986
Computer science faculty submit proposal to form separate Department of Computer Science

1986

DEC 1985
Computer Science Coordinating Committee is appointed.

OCT 1985
San Diego Supercomputer Center founded under Prof. Sid Karin with NSF grant, becoming part of Internet backbone

1985

1982
Division of Engineering created under Dean Lea Rudee

1982

1975
$35,000 approved expenditure for stand-alone minicomputers for use in the AP&IS introductory computer science courses

1979
AP&IS becomes Electrical Engineering & Computer Sciences Department

1979

1974
UCSD Pascal developed by Prof. Kenneth Bowles with 70 students

1974

1968
Applied Physics and Information Science (AP&IS) Department created with 17 faculty

1967

1967
Construction of AP&M Building, first home to the CSE Department

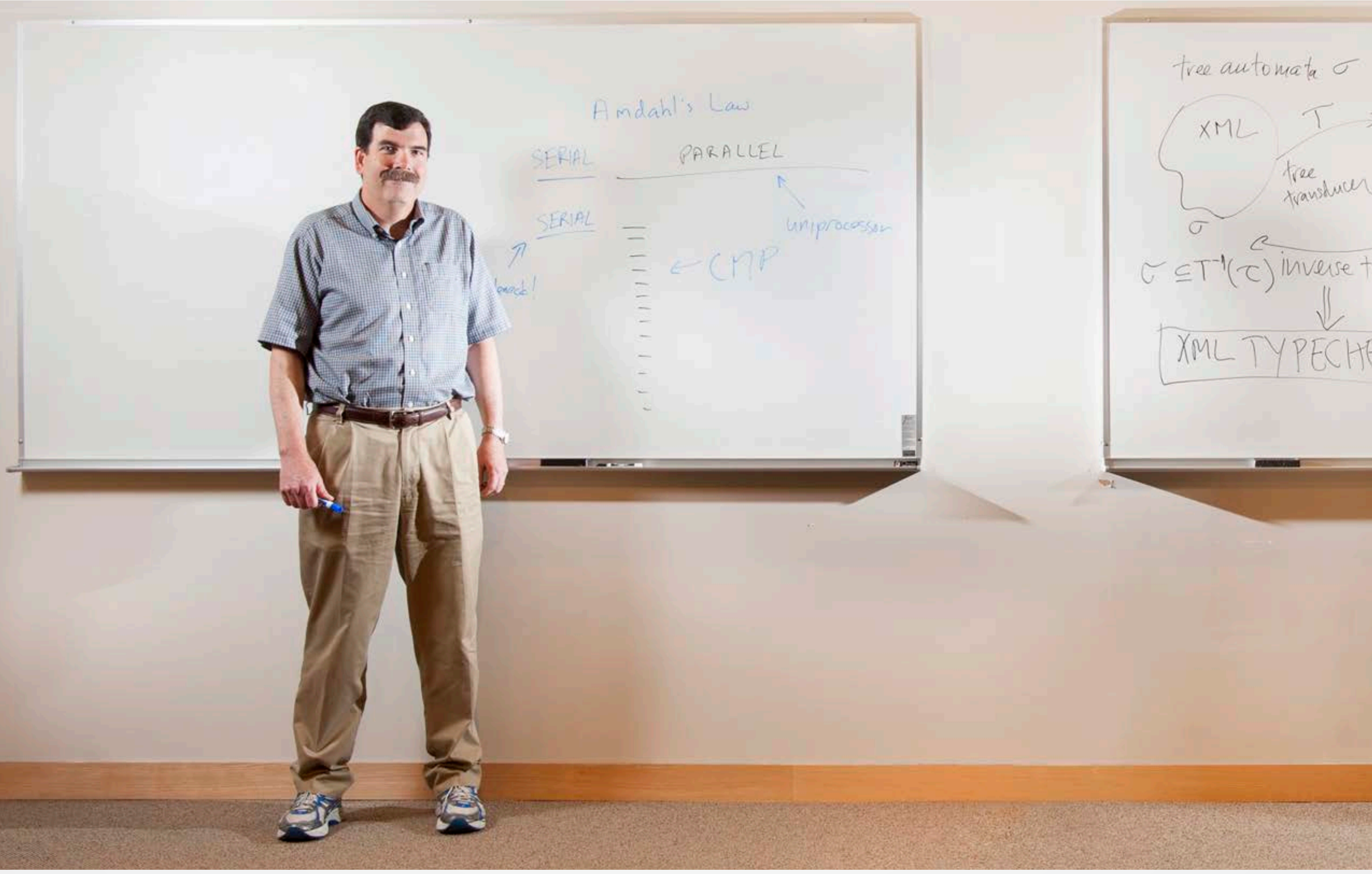
1965
Applied Electrophysics Department created

Standing the Test of Time

We all pursue our work in the hope that it will change the world, but lasting impact is only visible in retrospect. Several recent awards illustrate how well our research is standing this test of time.

Take the work of Prof. Dean Tullsen (picture below at left). In 2011, he (along with his co-authors) received the International Symposium on Computer Architecture (ISCA) Influential Paper Award. To determine the winner, ACM's Special Interest Group on Computer Architecture and the IEEE Computer Society Technical Committee on Computer Architecture looked back at the ISCA proceedings from 15 years earlier, and selected one paper that had the greatest influence in the computer architecture field in the intervening period. Tullsen's 1996 paper, "Exploiting Choice: Instruction Fetch and Issue on an Implementable Simultaneous Multithreading Processor," earned the 2011 award. However, even more impressive, is that this is not the first time Tullsen's work has been selected for this honor. Just one year earlier he won the same award for his 1995 paper on "Simultaneous Multithreading: Maximizing On-Chip Parallelism." Taken together, the papers laid a foundation for widespread commercial use of simultaneous multithreading in a range of settings, including by Intel under the Hyperthreading brand name.

Also in 2010, Prof. Victor Vianu (below right) was singled out for a paper published 10 years earlier in the Proceedings of the Symposium on Principles of Database Systems (PODS). The ACM PODS Alberto O. Mendelzon Test-of-Time Award honored the paper that had the most impact in terms of research, methodology or transfer to practice over the previous decade. Vianu shared the award with co-authors of "Typechecking for XML Transformers," a paper which determined that a key problem – checking whether or not an XML transformation is well typed – is indeed decidable. The paper is extensively cited in the literature and, as described by the award committee, had "a major influence on the methodology and direction of subsequent research on XML data modeling and management." In 2012, the American Association for the Advancement of Science (AAAS) inducted Prof. Vianu as Fellow in recognition of this body of work.



(L-R) Professors Dean Tullsen and Victor Vianu were cited for their lasting contributions in the fields of multithreading processors and XML data modeling, respectively.

ACM Fellows

The Association of Computing Machinery (ACM) is the world's largest educational and scientific computing society and each year it recognizes a few dozen members as Fellows to honor their "achievements in computer science at information technology". Fewer than one percent of all ACM members are ever named Fellows.

Yet between 2010 and 2012, ACM elected six new Fellows from UC San Diego – one in 2012, three in 2011, and two in 2010.

The most recent inductee, in December 2012, was Andrew Kahng, cited for his "contributions to physical design automation and to design for manufacturability of microelectronic systems." The previous year, professors Keith Marzullo, Dean Tullsen and Amin Vahdat were among only 46 researchers worldwide who were recognized for their contributions to computing. The awards specifically recognized the department's leadership in networking and computer architecture as well as distributed computer systems.

Keith Marzullo

Dean Tullsen

Amin Vahdat

Marzullo was cited for "contributions to distributed systems and service to the computing community," notably during his current leave of absence to serve as Division Director of the Computer and Network Systems Division at NSF headquarters. Tullsen was recognized for contributions to the architecture of high-performance processors. He co-directs the High Performance Processor Architecture and Compilation Lab. ACM honored Amin Vahdat for his contributions to datacenter scalability and management. When the ACM Fellowships were announced, Vahdat was on leave at Google, working on datacenter and wide-area network architectures.

Finally, two UC San Diego professors were named ACM Fellows from a class of 41 worldwide. professors Pavel Pevzner and Stefan Savage. Pevzner, a Howard Hughes Medical Institute professor, was

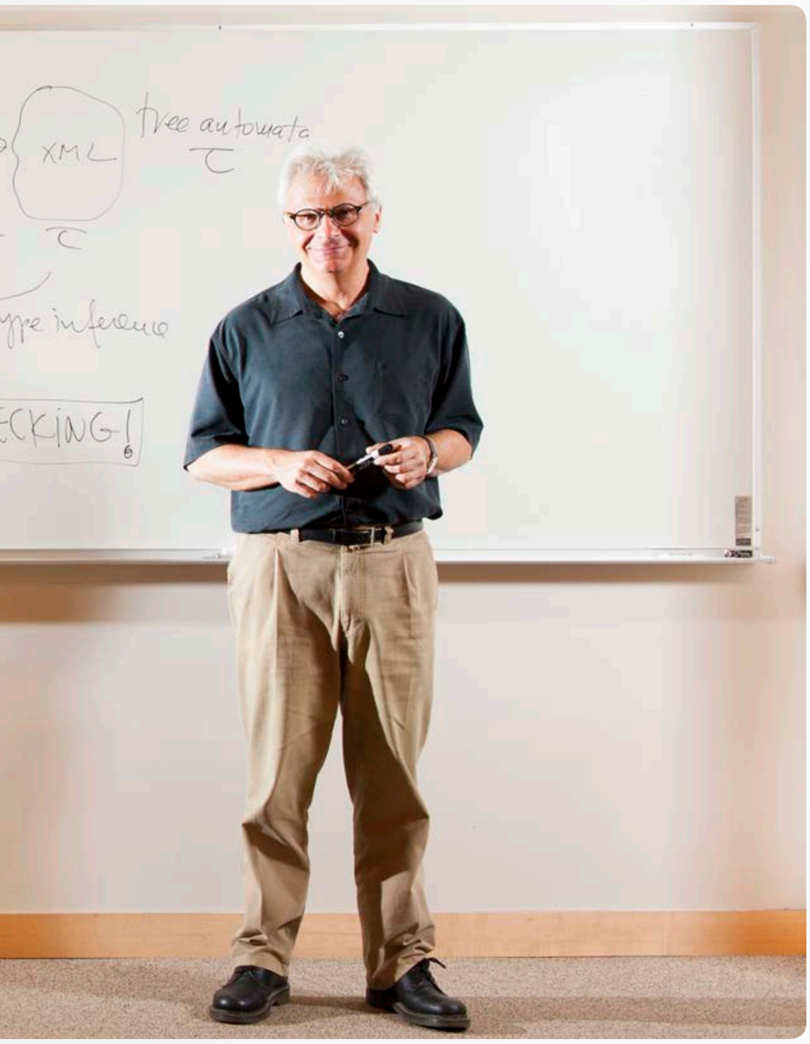
Defining the Semiconductor Roadmap

The Computer Science and Engineering department now boasts seven active faculty members who are Fellows of the Institute of Electrical and Electronics Engineers (IEEE).

In 2010 Andrew Kahng joined their ranks. IEEE cited his contributions to "the design for manufacturability of integrated circuits, and the technology roadmap of semiconductors." Indeed, the UC San Diego computer science alumnus (Ph.D. '89, M.S. '86) has for the past 15 years played a pivotal role in the development of the International Technology Roadmap for Semiconductors (ITRS), primarily guiding its work to predict future system drivers and design technology requirements. Kahng has also served on the editorial boards of IEEE Transactions on VLSI, IEEE Transactions on Circuits and Systems I, and IEEE Design and Test (where he contributes "The Road Ahead," a regular column focused on underlying technologies and future developments in the design and test of integrated circuits).

Kahng joined Dean Tullsen (2009), Rajesh Gupta (2004), Jeanne Ferrante (2005), Bill Howden (2001), C.K. Cheng (2000) and Walter Burkhard (2000) as IEEE Fellows. Former CSE faculty elected IEEE Fellows include Larry Carter (2000) and Andrew Chien (2007). IEEE is the world's largest professional association dedicated to advancing technological innovation and excellence for the benefit of humanity.

Andrew Kahng was elected a Fellow of IEEE in 2010, and a Fellow of ACM in 2012 (story starts at bottom of page four)



recognized for his contributions to algorithms for genome rearrangements, DNA sequencing and proteomics. His work has focused broadly on algorithmic aspects of bioinformatics, and interdisciplinary approaches to bioinformatics education.

ACM honored Stefan Savage for his "contributions to large-scale systems and network security." His work has focused on security threats enabled by broad Internet connectivity.

Other ACM Fellows on the UC San Diego faculty include Victor Vianu (2006), George Varghese (2002), Ron Graham (1999) and Jeanne Ferrante (1996).

Pavel Pevzner

Stefan Savage

Finally, a number of past faculty have been named Fellows while at UCSD, including Venkat Rangan (1998), Fran Berman (2000), Sid Karin (2000), George Varghese (2002) and Andrew Chien (2002).

Large-Scale Federal Awards

CSE professors lead large-scale, multi-institutional projects funded by federal agencies, including the National Science Foundation (NSF) and National Institutes of Health (NIH). Among them:

In September 2012, NSF awarded a $10 million, five-year Frontier grant to computer scientists at UC San Diego, the International Computer Science Institute (ICSI) at Berkeley and George Mason University to map out the illicit activities taking place in the cybersecurity underworld and to understand how the mind of a cybercriminal works. Prof. Stefan Savage is one of the lead researchers of the new **Center for Evidence-based Security Research** (CESR). "Fighting cyber threats requires more than just understanding technologies and the risks they're associated with," says Savage. "It requires understanding human nature." Other faculty on the project include Geoffrey M. Voelker, Alex Snoeren and Lawrence Saul. (For more on cybersecurity research in CSE, see pp. 20-21.)

Since 2006 Prof. Garrison Cottrell has directed the **Temporal Dynamics of Learning Center** (TDLC) to explore the role that time plays in learning. In 2011, the NSF renewed the TDLC grant to the tune of $18 million that will benefit a team of scientists and educators that includes more than 40 investigators from 17 partner research institutions, including UC Berkeley, Rutgers University at Newark, and Vanderbilt University. "Our team is working together to unravel the mysteries of learning through cooperative, interdisciplinary science," explains Cottrell. "People think that's hard to do, but at UC San Diego we thrive on this approach."

Prof. Rajesh Gupta led a consortium of six universities (UCSD, UCLA, UC Irvine, Stanford, University of Michigan and University of Illinois at Urbana-Champaign) to win a $10 million NSF Expeditions in Computing award in 2010 (see p. 6). The **Variability Expedition** proposes a new class of computing machines with a fluid software-hardware interface will mitigate the variability of manufactured systems and make machines robust, reliable, and responsive to changing operating conditions. CSE co-PIs on the Expedition project include Ranjit Jhala, Sorin Lerner, Tajana Simunic Rosing, Steven Swanson, and YY Zhou. "Changing the way software interacts with hardware," said Gupta, "offers the best hope for perpetuating the fundamental gains in computing performance... of the past 40 years."

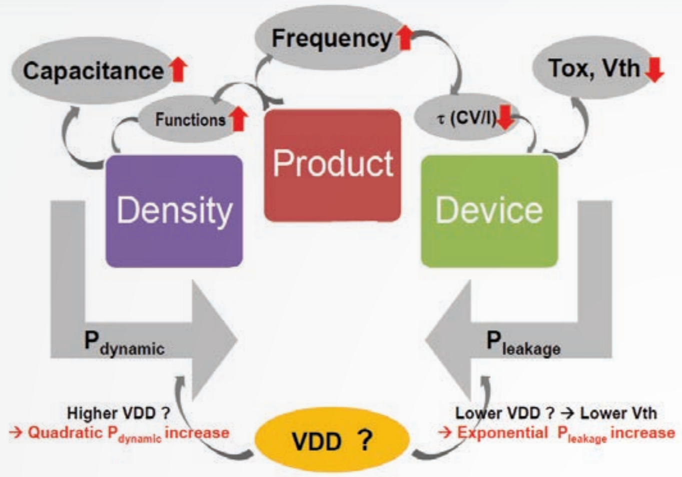
NIH awarded nearly $5 million over five years to a group of proteomics experts led by CSE Prof. Pavel Pevzner to fund a **Center for Computational Mass Spectrometry**. The 2008 grant funds development of algorithms and software for deciphering all the proteins present in biological samples to improve vaccine development, cancer diagnostics, and more.

Energy Efficiency and Resilience in Future ICs

Many energy efficiency gains worldwide have been due to computer science and engineering technology. So it should come as no surprise that CSE and ECE Prof. Andrew Kahng sees another power revolution ahead – thanks to processors designed to permit more processing with less energy. As a leader in the global effort to create a roadmap for the future of both semiconductor technology and its market drivers, the computer engineering professor is involved with a number of research projects that are rethinking the way processors are designed to enable higher performance using less power.

Kahng's group takes a broad view of energy issues as they relate to semiconductor scaling, density (more functions per chip), device performance, and the leakage of increasingly expensive energy as scaling continues.

The 'power management gap' (pictured at right) will have increased 100 times from 1995 to 2025 based on current growth, and more research into low-power techniques – including work at UCSD on power gating, transistor channel-length biasing, guardband reduction, and dynamic voltage/frequency scaling – will be critical to bridge the gap.



A 'power management gap' arises from conflicting product demands for IC performance (demanding high voltage) and energy efficiency (demanding low voltage).

In one recent project, working with a team from UC San Diego and the University of Illinois at Urbana-Champaign, Kahng has developed a novel approach to reducing the power consumption of processors by enabling tradeoffs between voltage and reliability.

"This new area of resilient design requires a different mindset," explains Kahng. "To further reduce energy per instruction while scaling voltage down, we must move from worst-case design to a better-than-worst case design, living with variations, and dynamically managing reliability (errors)."

Power has always been a primary design constraint for processors. When voltage is reduced, frequency follows suit, unless the user is willing to accept a drop in performance in the form of timing errors which occur because voltage reductions increase signal delays.

In their research, the scientists found that after scaling voltage by only 10 percent (and keeping frequency fixed), a circuit module completely breaks down. The error-rate curve is so steep at this 'critical operating point' that today's processors cannot even begin to trade reliability for extra power savings.

Kahng and his colleagues developed a new computer-aided design (CAD) now referred to as 'recovery-driven' design. Instead of designing microprocessors that use added energy to operate error-free, recovery-driven design allows occasional timing errors in the processor to be tolerated because of a built-in error-resilience mechanism.

The end result: 'underdesigned' chips that require significantly less power than required by conventional hardware that is, by definition, 'overdesigned'.

"There are critical paths that are more critical than others," said Kahng in a March 2012 journal article*.

"The power benefits of exploiting error resilience are maximized by redistributing timing slack from paths that cause very few errors to frequently exercised paths that have the potential to cause many errors, " he added. "Significant power benefits result when we deliberately allow errors to result from voltage overscaling, while relying on error resilience techniques to tolerate these errors."

Kahng and his coauthors reduce the error rate at a given voltage, and thereby reduce the minimum supply voltage and power for a target error rate. The target error rate is based on how many errors can be gainfully tolerated by a hardware or software error-resilience mechanism. "This has opened the door to further research at UCSD to reduce the guardband, or margin, used in the design process - and to even redesign basic arithmetic circuits to be accuracy- and energy-tunable," Kahng said.

Significant energy gains appear possible from recovery-driven design: in their study, Kahng's team observed an average 27 percent power reduction overall. Will these and other energy-reliability tradeoffs yield significant power savings in the future? Kahng is certain of it.

* **Recovery-Driven Design: Exploiting Error Resilience in Design of Energy-Efficient Processors**, A.B. Kahng, S. Kang, R. Kumar and J. Sartori, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 31, No. 3, March 2012, pp. 404-417. http://bit.ly/SFklGc

Energy Efficiency and Variability

Manufactured parts can differ in power consumption by as much as 10X from one part to the next when taking into account their active and sleep power profiles. The Variability Expedition (variability.org) seeks to build a software stack for a class of Underdesigned and Opportunistic Computing (UNO) machines that purposely expose such variations to the system stack, including software. Working at the intersection of software adaptability and hardware error resilience, the envisioned UNO machines forge a new path for relaxed IC design implementation and manufacturing constraints. "Such machines will enable new computing models that admit computation on relaxed reliability parts or relaxed acceptance criteria for computational results ," explains the project PI, Prof. Rajesh Gupta.

Energy Efficiency Through Virtual Machine Management

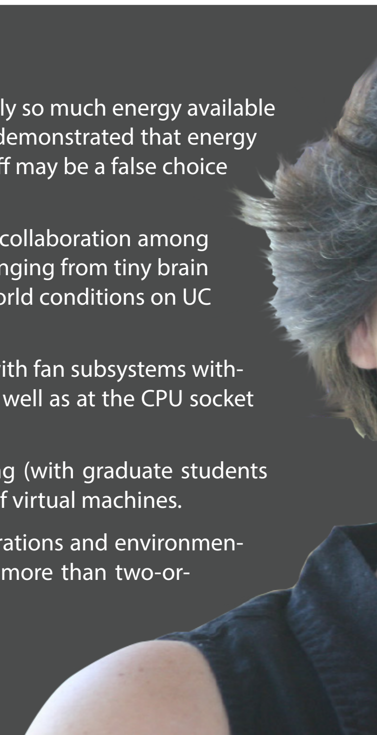
Until now, making large-scale datacenters more energy efficient has meant sacrificing performance, since there is only so much energy available in a system to handle all computing operations. But a series of projects co-led by Prof. Tajana Simunic Rosing have demonstrated that energy costs can be reduced across scales, while simultaneously boosting performance – proving that the historical trade-off may be a false choice in many cases.

Rosing is the principal investigator for the Large-Scale Systems Thrust of the Multi-Scale Systems Center (MuSyC), a collaboration among UC San Diego and nine other universities charged with finding ways to improve the design of computing systems ranging from tiny brain sensors to large datacenters. Her System Energy Efficiency (SEE) Lab tests new, energy-saving methods under real-world conditions on UC San Diego's NSF-funded GreenLight Instrument testbed.

In a 2011 journal article*, Rosing proposed a multitier approach to lower substantially the cooling costs associated with fan subsystems without compromising the system's performance – doing so by allocating the workload intelligently at the core level as well as at the CPU socket level.

That paper followed a 2010 article in *ACM Transactions of Design Automation of Electronic Systems,* in which Rosing (with graduate students Gaurav Dhiman and Giacomo Marchetti) laid out what she calls 'vGreen,' a system for energy-efficient management of virtual machines.

By harnessing what's known as 'virtual machine management' – a method for remotely monitoring datacenter operations and environmental variables, such as temperature, inside server rooms – Rosing and her colleagues have been able to achieve a more than two-orders-of-magnitude savings in energy costs, and in some cases increased performance.



Smart Buildings, Saving Energy

CSE Research Scientist Yuvraj Agarwal received UCSD's Outstanding Faculty Award for Sustainability for 2012 – fitting recognition for a series of research projects and novel solutions that are helping to make offices and buildings more energy efficient. One example: the UC San Diego Energy Dashboard, a Web portal that allows anyone to track how much power is being used by any building on the UC San Diego campus (and even floor by floor in the CSE Building).

Since then, Agarwal and Ph.D. student Thomas Weng have published a groundbreaking paper in the journal *IEEE Design and Test Special Issue on Green Buidings** on the use of sensing and actuation to improve energy efficiency.

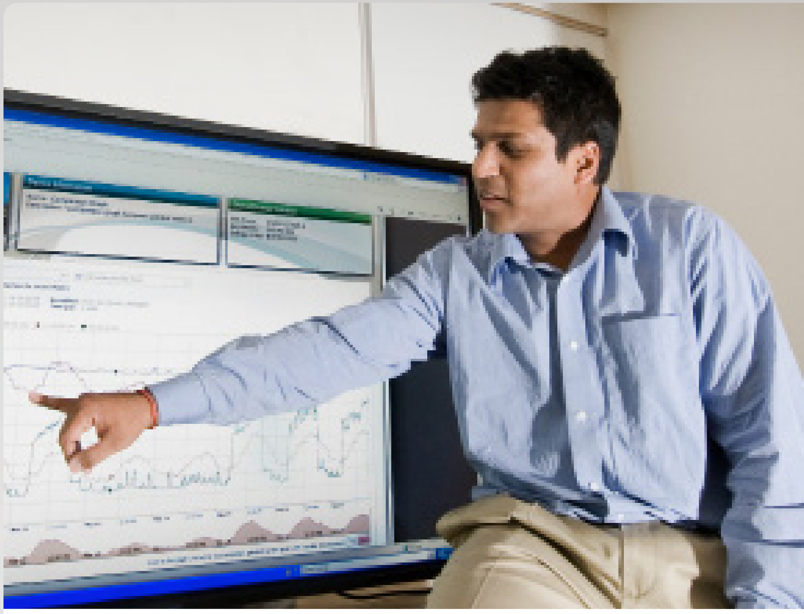
The researchers used CSE's headquarters building as a stand-in for mixed-use commercial buildings, which account for roughly 35 percent of total electricity consumed in the U.S. each year. "Even a small improvement in energy efficiency could translate into substantial savings here and around the world," says Agarwal.

To understand where energy is wasted, Agarwal and Weng looked at the major systems which use power, day in and day out: heating, ventilation and air-conditioning (HVAC); lighting; information technology (IT) equipment; and various devices that occupants plug into office outlets (so-called 'plug-load' devices such as lights and fans).

Actuation is critical. Agarwal argues that the smart building of the future must be able to turn off most devices when occupants are not using a space – an idea he calls 'aggressive duty-cycling' of buildings. Sensors are needed to track the presence of occupants, and they must be tied into actuation systems that are 'smart' enough to turn a system off when it is not in use.

In addition to occupancy, building sensors can monitor environmental conditions and energy use itself. All three are important to monitor the status of the building 24/7 in real time.

Agarwal and his team had previously designed an occupancy sensor system that improves on standard passive-infrared (PIR) sensors (which may turn on the light if it detects someone entering an area). His group was able to produce a PIR sensor inexpensively but they also made it battery powered and wireless for easy deployment, and linked it to a contact switch that monitors whether the door is open or shut. If it is open, the system assumes someone is in the office, even if they aren't moving around. Accuracy improves further if a computer in the room is being used (for which Agarwal's team developed monitoring software tied into his Sleep-Server project).



Research scientist Yuvraj Agarwal points to a real-time chart of energy use in the CSE building as depicted on the UC San Diego Energy Dashboard, available at *http://energy.ucsd.edu.*

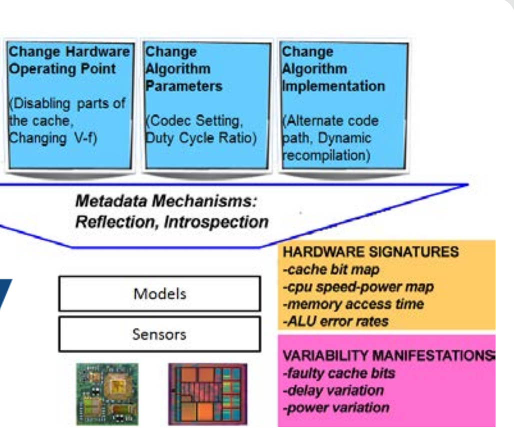
The researchers tested a novel control system of aggressive duty-cycling for HVAC on the second floor of the CSE building on one of two warm days. On the first, they used the building's standard HVAC controls; on the second day, they used real-time occupancy information from each sensor node to turn the floor's HVAC systems on or off. The result: the entire CSE building used more than 11 percent *less* energy – even though the control system was deployed only on one floor of the four-story building. Agarwal estimates that the CSE building could reduce its HVAC-related energy use by 30 percent if the smart-control system were deployed across the entire building.

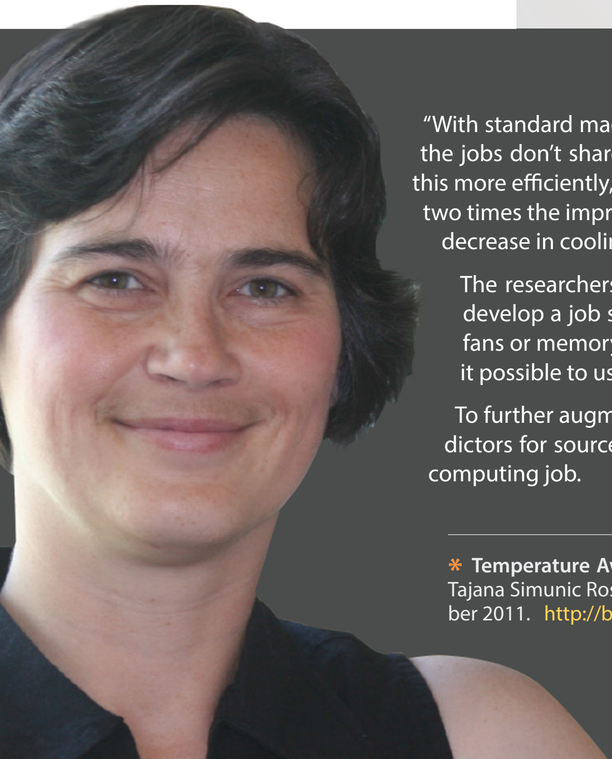
The study was carried out with the help of Ph.D. students Thomas Weng and Bharathan Balaji, and M.S. student Seemanta Dutta. Agarwal's Systems, Networking and Energy Efficiency (SYNERGY) lab also developed its own Synergy Smart Energy Meter, a plug-load meter that can be used to disconnect the electricity and shut off a device, and it can be operated remotely over a wireless network.

As for the cost of installing all of these sensors and actuation systems, Agarwal estimates the total cost to be approximately the same as the yearly energy savings at 13 cents per kilowatt-hour.

"That means a building can recover the installation costs in one year through the reduction in energy usage," he says. "While this will be different for every building, smart building systems are now extremely economical and provide an excellent return on investment for buildings where they are deployed. Of course, being green by helping the environment is also a definite plus!"

* **From Buildings to Smart Buildings – Sensing and Actuation to Improve Energy Efficiency**, Thomas Weng and Yuvraj Agarwal, *IEEE Design and Test, Special Issue on Green Buildings*, July-August 2012. http://bit.ly/Xkjw8i


Change Hardware Operating Point
(Disabling parts of the cache, Changing V-f)
Change Algorithm Parameters
(Codec Setting, Duty Cycle Ratio)
Change Algorithm Implementation
(Alternate code path, Dynamic recompilation)
Metadata Mechanisms: Reflection, Introspection
HARDWARE SIGNATURES
-cache bit map
-cpu speed-power map
-memory access time
-ALU error rates
Models
Sensors
VARIABILITY MANIFESTATIONS
-faulty cache bits
-delay variation
-power variation



"With standard machine management, when you mix two types of computing jobs, the quality suffers because the jobs don't share energy resources in a fair way," explains Rosing. "What we've developed is a system to do this more efficiently, and to more efficiently cool the subsystem along with the job scheduling. As a result, we see two times the improvement in energy efficiency in terms of the power needed to run the jobs, and a 70 percent decrease in cooling cost, with excellent performance."

The researchers have designed the system at the software level by using the inputs of various sensors to develop a job schedule and control methodology for setting specific parts of a subsystem (such as cooling fans or memory) at appropriate levels. A related system of battery-driven, distributed energy storage makes it possible to use stored energy during times of peak power usage.

To further augment the energy savings provided by these systems, the researchers have also developed predictors for sources of green energy, such as solar and wind power, to determine the best time to schedule a computing job.

* **Temperature Aware Dynamic Workload Scheduling in Multisocket CPU Servers**, Raid Ayoub, Krishnam Indukuri and Tajana Simunic Rosing, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 30, No. 9, September 2011. http://bit.ly/NSTB17

Mining – and Minding – Her Ps and Qs

Each time you connect to a secure Web site (say a bank's site), you begin by downloading a certificate published by the site, which asserts that its Web address is legitimate and contains a public key that your computer can use to establish a secure connection. The public key, ostensibly, prevents anyone else from spying on your connection.

But according to a paper presented at the 21st USENIX Security Symposium in August 2012, vulnerable public keys are "surprisingly widespread" on the Internet, especially for certain types of devices such as routers and firewalls. The paper*, "Mining Your Ps and Qs: Detection of Widespread Weak Keys in Network Devices," won the conference's Best Paper award.

UC San Diego postdoctoral researcher Nadia Heninger co-authored the paper with three colleagues from the University of Michigan: Zakir Durumeric, Eric Wustrow and J. Alex Halderman. To pursue their research, Heninger and her colleagues scanned the entire Internet in 24 hours and collected public keys from 22 million hosts, which were using them to secure Web and SSH connections.

The researchers found evidence that public keys for hundreds of thousands of devices were insecure because they had been generated in a way that would allow anyone to easily calculate the private keys. Furthermore, devices from dozens of manufacturers – 54 cited in the paper – proved vulnerable, and the researchers informed all of them prior to publishing their results.

Two cryptographic algorithms have been the *de facto* standards used for these public keys: RSA, an acronym that derives from the last names of inventors Ronald Rivest, Adi Shamir, and Leonard Adleman; and DSA, the U.S. federal standard Digital Signature Algorithm.

The problem, says Heninger, is that some of these public keys are not sufficiently random. "These public-key algorithms are supposed to be designed so that it is impossible for someone to figure out the private key just by looking at a public key," she explains. "But because these keys were not truly random, we were able to use mathematical relationships between pairs of keys to calculate their private keys."

If two different devices have the same public key, they also have the same private key, which means that malicious users could gain access to restricted content in one location if they merely decode the public key for the other.

Heninger says her team was able "to remotely compromise about 0.4 percent of all the public keys used for SSL [Secure Socket Layer] Web site security." The SSL 'handshake' protocol typically uses RSA encryption, which consists of two numbers – one of which is the product of two randomly chosen prime numbers, *p* and *q*, which are generated by the RSA key-generation algorithm.

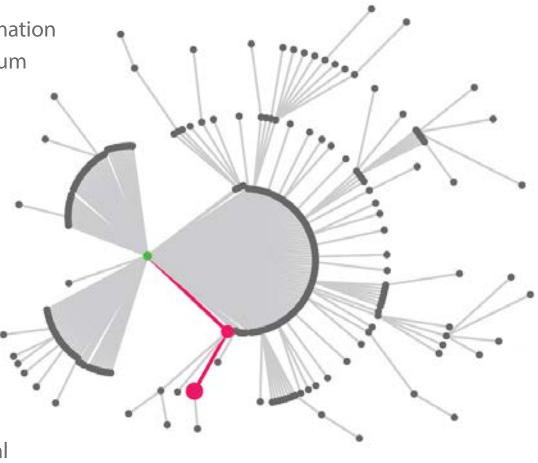
Fortunately, servers and most large Web sites were in the clear: all of the compromised keys were "for various kinds of routers and firewalls and VPN servers – no banks," says Heninger. These 'unsigned' certificates had been automatically generated by 'headless' devices and were not sufficiently random, whereas the vast majority of certificates that were signed by a certificate authority (and most likely had been generated by humans) appeared secure.

The only fix, according to Heninger, is for device manufacturers and software developers to "make sure they generate their keys with good randomness." She and her colleagues have developed an online service that lists all of the compromised keys they discovered, so users can check keys against them.

"This is a wake-up call to the security community," concludes Heninger. "It's a reminder to all of how security vulnerabilities can sometimes be hiding in plain sight."

* **Mining Your Ps and Qs: Detection of Widespread Weak Keys in Network Devices**, Nadia Heninger, Zakir Durumeric, Eric Wustrow, and J. Alex Halderman, *Proc. 21st USENIX Security Symposium*, August 2012. http://bit.ly/Nxi12e

The Challenge of Network Science



Overview of the different levels of a crowd recruited by Cebrian's team in recent challenges

As a research scientist in CSE, Manuel Cebrian studied network science, and in particular, how information spreads across large platforms such as social networks. The field is so new that Cebrian turned to a new forum to understand the phenomenon of crowdsourcing: competitive challenges.

When he was still a postdoctoral researcher at MIT, Cebrian was on the team that entered and won the DARPA Network Challenge, which required teams to locate 10 red weather balloons dotted around the United States. Using what was then a novel tool – crowdsourcing – Cebrian and his colleagues took barely nine hours to find all of the balloons. They did so by enlisting social networks on a grand scale and by learning how to filter out misinformation. For its part, DARPA was seeking to explore "the roles the Internet and social networking play in the timely communication, wide-area team-building, and urgent mobilization required to solve broad-scope, time-critical problems."*

After moving to UC San Diego, Cebrian put together a team in late 2011 to enter the Department of Defense's Shredder Challenge, in which he faced another type of adversity. The challenge was to solve a puzzle by patching together 10,000 pieces from several shredded documents. Cebrian turned to social networks and the media to recruit more than 3,600 people to solve the puzzle, but half-way through the challenge, disaster struck. The CSE team was among the top three contenders when their Web site was hit by a series of sophisticated attacks. Over three days, the attacker removed pieces that had been successfully matched to neigh-

Mihir Bellare works in cryptography and security, particularly practical, proven-secure schemes. He co-designed HMAC, used in SSL, TLS, SSH, IPSEC. He received the ACM Kanellakis Theory and Practice Award, and RSA Conference Award in Mathematics.

Fan Chung Graham is a professor in CSE and Mathematics, and holds the Paul Erdös Chair in combinatorics. Her other interests include spectral graph theory, algorithmic design and Internet computing. She is a Fellow of the American Academy of Arts and Sciences.

Ron Graham's research covers many areas in math and theoretical computer science, including combinatorics, graph theory, number theory, geometry, and the design and analysis of algorithms. He is a past president of the Int'l Jugglers Association.

Russell Impagliazzo has worked in complexity theory, cryptography, and algorithms. He studies the power of randomness and pseudo-random generators in computation and in cryptography, as well as optimization heuristics.

Lattice Cryptography, Trapdoors and the Cloud

With the explosive growth in confidential data housed in 'the cloud,' where it may seem more vulnerable than if it were located on a local hard drive, security experts are looking to cryptographers for help.

"Imagine running your most computationally intensive programs on your large data sets on a cluster of remote computers, and in a cloud computing environment, while keeping your programs, data, and results encrypted and confidential," wrote CSE Prof. Daniele Micciancio in a 2010 article in the *Communications of the ACM*.

The idea of such fully homomorphic encryption traces back to the earliest public-key encryption schemes in the 1970s, but it wasn't until 2009 that IBM's Craig Gentry proposed a new approach to constructing fully homomorphic cryptographic systems. "Now, most cryptographers (me among them) are convinced the Holy Grail exists," said Micciancio. "In fact, there must be several of them, more or less efficient ones, all out there waiting to be discovered."

Gentry's work was the first real breakthrough in using lattices to reach that Holy Grail. "Since then, a dozen other schemes have been published, and all are based on lattices," explains Micciancio, a theoretical computer scientist. "So from what we can tell, lattices are the only way to build homomorphic encryption, perhaps because lattices can mathematically support both addition and multiplication, so you can perform just about any operation."

Lattices have been used historically in math and coding theory, and many applications in computer science date back to the early 1980s. "They were even used in cryptoanalysis to break codes," notes Micciancio. "But it wasn't until the late 1990s that we started to find ways to use lattices to build secure cryptographic systems."

For his part, Micciancio set about proving the hardness of lattice-based codes – to prove that they couldn't be broken. "These days cryptography and security are everywhere on the Internet and you cannot wait a few years to see if a function works or not," he explains. "There is a strong theoretical basis and evidence that these functions are hard to break."

Reaching the Holy Grail, however, requires much greater efficiency in lattice cryptography in order to make it run faster while taking up less space. "My work is at the foundation of lattice cryptography," explains Micciancio. "Most of this foundational work will stay fresh for a long time. My current work is focused on trying to move lattice cryptography from theory to something that is efficient enough to be usable in practice, and that requires not just an engineering effort, not just implementation, but also the math that needs to be developed to achieve its maximum potential efficiency."

To that end, Micciancio delivered a paper* on the use of 'trapdoors' with lattices at the April 2012 Eurocrypt conference in Cambridge, England. Trapdoors have been used in cryptography since the term was coined in the mid-1970s. A trapdoor includes special information without which it would be impossible to do a computation in one direction, even though it is easy to compute in the other direction.

Trapdoors are fundamental building blocks used in many lattice cryptography applications. "The potential applications to securing the cloud are clear, though it is still far from being usable in practice," admits Micciancio. "The Eurocrypt paper is about generating random lattices together with secret trapdoor information that allows you to easily solve lattice problems that otherwise are cryptographically hard to solve."

Micciancio and co-author Chris Peikert have devised what they call 'strong trapdoors' that are simultaneously simple, efficient, easy to implement (even in parallel), and asymptotically optimal to get around a major bottleneck: the difficulty in performing a long series of tasks, one after the other. "Our techniques substantially improve on prior methods, both in terms of practical performance and the quality of the produced outputs," says Micciancio.

"The simple structure of the new trapdoor and associated algorithms can be exposed in applications, leading to further simplifications and efficiency improvements."

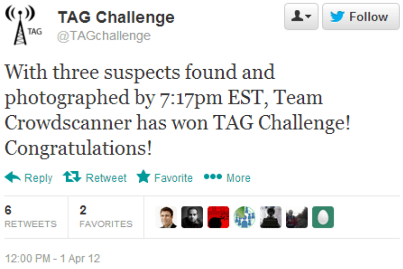
The computer scientist adds that more compact and efficient trapdoors appear necessary for bringing advanced lattice-based schemes into practice – in part because of unsatisfactory runtimes. In more recent work, Micciancio has come up with new, improved algorithms that reduce the worst-case running time for all lattice problems. The 'deterministic, single exponential time algorithm' breaks a speed barrier of sorts – and could have applications to both security and wireless communications.

* **Trapdoors for Lattices: Simpler, Tighter, Faster, Smaller,** Daniele Micciancio and Chris Peikert, *Eurocrypt 2012*, April 2012. http://bit.ly/PmgEYn

boring pieces, and pushed pieces off the UCSD team's puzzle board entirely. The setback forced the team to rewrite code to protect against such attacks, losing precious time and taking some of the enthusiasm out of 'the crowd.' In the end, the team placed sixth – and Cebrian stoically believes the sabotage may trigger a new field of network science, for which he coined the term, 'competitive crowdsourcing.'

Cebrian paid special attention to security when he joined forces with colleagues in Abu Dhabi and the U.K. to take part in the Tag Challenge, sponsored by the Department of State, in April 2012. The goal: to locate five fake jewel thieves in five cities and four countries within 12 hours. Using only a mugshot and brief story about each 'suspect,' Team Crowdscanner developed an incentive structure for dividing the prize money, rewarding not just people who found the fake thieves, but also anyone who recruited the finder, and one dollar for every person a participant recruited. Cebrian and his colleagues alerted the media and used Twitter and Facebook to spread the word. In the end, Team Crowdscanner was able to find the suspects in New York City, Washington D.C. and Bratislava, Czechoslovakia. That was good enough for Cebrian's group to take the top prize in the Tag Challenge, which they announced on Twitter (at right).

Interviewed in *Popular Science* magazine, Cebrian – who is now a researcher at the University of Melbourne – was asked what was learned from the Tag Challenge. "We thought this was impossible," he answered. "We really thought this was beyond what social networks can achieve. So this sets new limits on social media's capabilities. If the balloon challenge was the former record, this is definitely the new one."



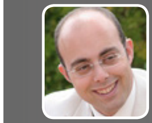
TAG Challenge
@TAGchallenge

With three suspects found and photographed by 7:17pm EST, Team Crowdscanner has won TAG Challenge! Congratulations!

6 RETWEETS | 2 FAVORITES

12:00 PM - 1 Apr 12

* **Time-Critical Social Mobilization,** Galen Pickard, Wei Pan, Iyad Rahwan, Manuel Cebrian, Riley Crane, Anmol Madan and Alex Pentland, *Science*, Vol. 334, No. 6055, pp. 509-512, October 2011. http://bit.ly/vTCU19

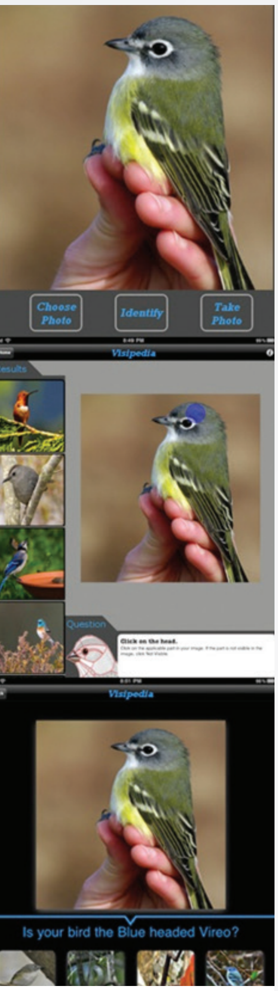


Shachar Lovett joined CSE in 2012 from the Institute of Advanced Study in Princeton, NJ. He works in the general field of computational complexity. His work in additive combinatorics led to a better understanding of algebraic property testing. He received his Ph.D. from Israel's Weizmann Institute of Science in 2010.

Daniele Micciancio works at the intersection of cryptography, computational complexity, and algorithms. His main research focus is lattice-based cryptography, an area of mathematical cryptography with the potential to deliver very fast and parallelizable cryptographic functions with strong security properties.

Mohan Paturi joined UCSD in 1986. He studies the theoretical underpinnings of computer science, efficient algorithms and their complexity. Paturi is also an expert in digital libraries, ontologies and data mining.

Hovav Shacham's research interests are in applied cryptography, systems security, and tech policy. He joined CSE in 2007, after participating in Secretary of State Debra Bowen's Top-to-Bottom review of the voting machines certified for use in California.



Visipedia app asks questions to help the user identify bird species.

Better Image Searches... Through Birding?

Searching text via Google or Wikipedia is a breeze compared to searching for just the right image on the Web. Now a team led by CSE Prof. Serge Belongie is determined to combine expertise in computer vision with citizen science in order to come up with a unique approach to image search. It's called Visipedia, and the initial proof of concept involves an ambitious plan to build a visually searchable database of over 500 North American bird species, curated by the Cornell Lab of Ornithology.

Computer scientists from UC San Diego, UC Berkeley and Caltech have designed Visipedia as a search system that can interact with users to provide more accurate results. The ultimate goal is to fill the image gap in the world of online search.

Belongie's group has developed an iPad app that will identify most birds – with a little help from birding enthusiasts. The app is essentially an interactive field guide, where the user submits an image, a computer-vision algorithm analyzes it, and if it cannot find a perfect match, Visipedia asks the user questions, e.g., "What color is the bird's breast?," until it can decipher a likely match.

The Visipedia team picked birds as a first test case for several reasons. "There is an abundance of excellent photos of birds on the Internet," explains Belongie. "The diversity in appearance across different species also presents a deep technical challenge, and most importantly, there is a large community of passionate birders who can put our system to the test."

The Cornell Lab is reaching out to those birders through its *http://allaboutbirds.org* Web site, and it hopes eventually to have 100 images identified by experts for each of 500-plus species. For their part, birders are eager to help train the Visipedia system because they know it will ultimately become a tool they can use in the field.

Users will be asked to upload their own images and if Visipedia cannot make a perfect match from what's in the database, it will prompt the user to describe specific details. Typically, the app takes about five or six questions to get it right, but as more birders use the app, the success ratio should improve.

Belongie hopes the app will also ignite a movement among other enthusiasts, who will partner with computer scientists to create similar apps for everything from flowers to butterflies or even mushrooms.

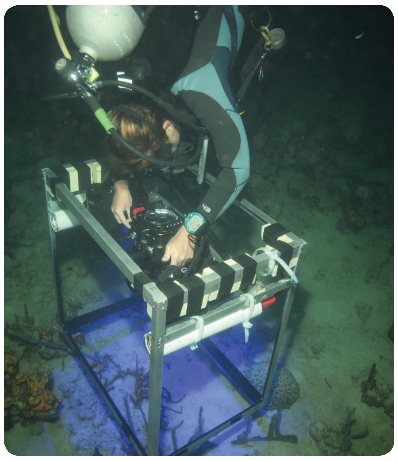
"The Visipedia architecture lends itself to the task of interactive, fine-grained categorization of highly general objects," explains Belongie. "We can't solve all the challenges by ourselves, but we believe our framework can."

A demo version of the user-friendly, mobile search app is available at *http://visipedia.org*. The user interface was designed by an undergraduate, while graduate students did much of the work on the search system itself.

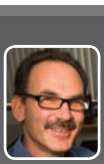
Computer Vision for Coral Reef Assessment

Across large areas of the world, coral reefs are dying from the twin effects of climate change and pollution. Scientists are already gathering enormous amounts of data in the form of images captured by underwater digital cameras and automated acquisition systems, but until now, the process of analyzing those images has been done by hand, one image at a time, which is both time-consuming and prone to error.

Enter a CSE computer-vision team, led by Prof. David Kriegman. Using computer-vision techniques, Kriegman and his team have developed a method of automatic annotation of coral reef images. In a 2012 paper*, working with colleagues from the Scripps Institution of Oceanography and Cal State Northridge, Kriegman showed that the new method can be 83% accurate, and went on to propose a benchmark dataset. The authors also showed that their new method accurately estimates coral coverage across reef sites over multiple years. Says Kriegman: "We think this offers exciting potential for large-scale coral reef analysis."

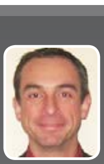


* **Automated Annotation of Coral Reef Survey Images**, Oscar Beijbom, Peter J. Edmunds, David I. Kline, B. Greg Mitchell, and David Kriegman, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012. http://bit.ly/RNwMSm



David Kriegman's research in computer vision includes 3D scene reconstruction, illumination and reflectance modeling, and object recognition with application to face recognition, microscopy, underwater imaging, robotics, and computer graphics.

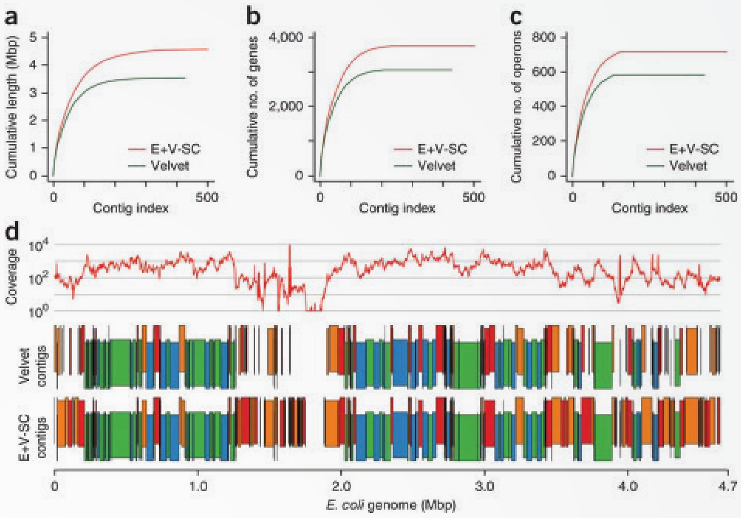
Serge Belongie works in computer vision and machine learning. His projects include human- in-the-loop object recognition for fine grained visual categories (Visipedia) and assistive technology for the visually impaired.



Henrik Wann Jensen develops computer graphics algorithms. He is best known for inventing the photon mapping algorithm for simulating global illumination, and the first practical technique for simulating translucent materials such as human skin.

Sequencing the Dark Matter of Life

Entire species of bacteria have traditionally been off-limits when it comes to DNA sequencing, because they cannot be cultured to provide the roughly one billion identical cells required by standard sequencing methods. This so-called 'dark matter' of life includes the lion's share of bacterial species living on the planet, including microorganisms that could yield new antibiotics and biofuels, and microbes living in the human body. "This part of life was completely inaccessible at the genomic level," says CSE Prof. Pavel Pevzner, a pioneer of algorithms for modern DNA sequencing technology.



Comparison of the popular open-source assembly program Velvet, and the EULER+Velvet-SC algorithm for single-cell assembly, for *E. coli,* as described in the Nature Biotechnology paper

Now a team of scientists including Pevzner, graduate student Hamidreza Chitsaz, and UCSD Mathematics Prof. Glenn Tesler (a former postdoctoral researcher in CSE), have developed an algorithm that dramatically improves the performance of software used to sequence DNA produced from a single bacterial cell. The breakthrough allows researchers to assemble virtually complete genomes from DNA extracted from a single bacterial cell. Using the new algorithm to enhance Multiple Displacement Amplification (MDA) technology developed a decade ago, scientists can start with just one cell, but capture 90 percent of genes – even for species which could not be sequenced until now.

If that sounds like magic, the new algorithm comes close: the innovation not only makes it possible to sequence the DNA of most bacterial life, but also to do so economically. The ramifications for human health could be astounding – especially given that bacteria account for roughly 10 percent of the weight of the human body.

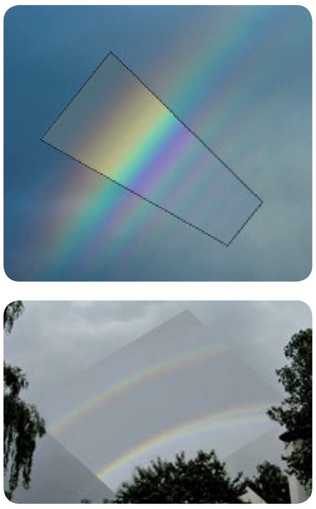
As described in the journal *Nature Biotechnology*, the scientists from UCSD, J. Craig Venter Institute (JCVI) and Illumina, Inc., noted that a vast majority of bacteria — unlike *E. coli*, which has been sequenced — cannot be cultured in the lab because they live in specific conditions and environments that are hard to reproduce, e.g., in symbiosis with other bacteria or on an animal's skin.

Since 2005, study co-author Roger Lasken (now at JCVI) – who developed MDA – has used the technology to sequence DNA produced from a single cell. MDA makes copies of that cell, but with errors that are not amplified uniformly. Modern sequencing algorithms are not equipped to deal with these disparities: they tend to discard bits of the genome that were replicated only a few times as sequencing errors, even though they could be key to sequencing the whole genome. The algorithm developed by Pevzner's team changes that. It retains these genome pieces and uses them to improve sequencing.

As a demonstration, Pevzner and his colleagues analyzed a species of marine bacteria never before sequenced, and were not only able to reconstruct its genome but also used the resulting sequence data to infer how the bacterium moves and produces energy. This genetic sequence is now available to researchers through the National Institutes of Health's GenBank database and, through this technique, it promises to be the first of many more.

* **Efficient de novo assembly of single-cell bacterial genomes from short-read data sets**, Hamidreza Chitsaz, Glenn Tesler, Pavel A. Pevzner and Roger Lasken et al., *Nature Biotechnology*, Volume 29, pp. 915-921, September 2011. http://bit.ly/oTLJdV

GRAPHICS + VISION (continued)



Simulations (inserts) of multiple supernumerary bows (top), and rare twinned rainbow (below)

Somewhere Over the (Simulated) Rainbow

The field of computer graphics is ever trying to better reproduce the imagery of nature, but sometimes this requires advancing our understanding of physics as well. For example… rainbows. The most accurate method for simulating rainbows, Lorenz-Mie theory, captures the effects of dispersion, polarization, interference and diffraction, but it still cannot replicate a variety of naturally occurring phenomena – such as double rainbows. The reason, it turns out, is the shape of the raindrops. In 2012 CSE Prof. Henrik Wann Jensen – who shared in an Academy Award in 2004 for an algorithm that allowed special-effects wizards to create life-like skin tones – published* the first comprehensive model of rainbows suitable for applications in computer graphics. The technique simulates the interaction of a wavefront of light with a physically-based shape of the water drop. The model matches Lorenz-Mie theory for spherical particles, but it also enables the accurate simulation of non-spherical particles. "Our technique is based on ray tracing extended to account for dispersion, polarization, interference and diffraction," explains Jensen. "We show how the non-spherical raindrops influence the shape of rainbows, and we provide a simulation of the rare twinned rainbow, which is believed to be caused by non-spherical water drops." The result: a graphics technique that can simulate many different rainbow phenomena.

* **Physically-Based Simulation of Rainbows**, Iman Sadeghi, Adolfo Munoz, Philip Laven, Wojciech Jarosz, Francisco Seron and Diego Gutierrez, and Henrik Wann Jensen, *ACM Transactions on Graphics*, Volume 31, Number 1, Article 3, January 2012. http://bit.ly/NtOyGB

Vineet Bafna works in computational biology, including computational proteomics, algorithms for genetics, and bioinformatics. Recent work focused on algorithmic problems in genetics, computational mass spectrometry, and comparative genomics.

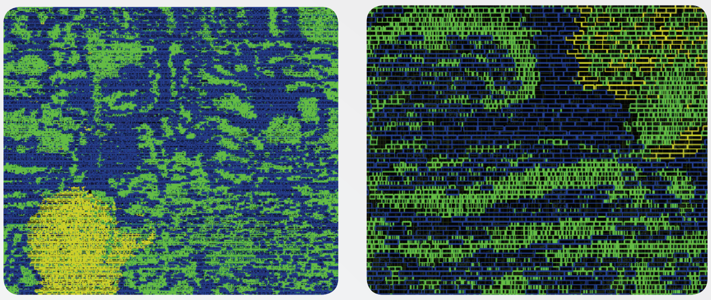
Pavel Pevzner studies computational molecular biology and bioinformatics. His recent work involves analysis of genomic rearrangements and identification of evolutionary 'faults' in human genome, sequencing new antibiotics, and single cell genomics.

GreenDroid vs. Dark Silicon

Usually microprocessors are designed and software developers then design apps to run on them. But a team led by CSE Prof. Michael Bedford Taylor, who now leads the UCSD Center for Dark Silicon, took a different approach: Why not use information about the most-popular mobile apps – music streaming, video, GPS, email and so on – to customize the design of a mobile application processor that will handle those apps more efficiently?

If the GreenDroid concept gains traction among semiconductor makers, it could revolutionize smart phones in the next five to 10 years. Based on simulations, a GreenDroid chip would be 11 times more efficient that the general-purpose processors in today's smart phones. For consumers, that means being able to use their phones much longer before needing to recharge the battery.

While extending the charge of smart-phone batteries would be a major benefit of GreenDroid chips, the project also tackles a looming problem for the industry: dark silicon. As manufacturers pack more and more functions onto microprocessors, "large swaths of a chip's silicon area… must remain mostly passive in order to stay within the chip's power budget," wrote Taylor and CSE Prof. Steven Swanson in *IEEE Communications**, adding that "only 1 percent or so of a modest-sized 32 nm mobile chip can switch at full frequency within a 3 W power budget."

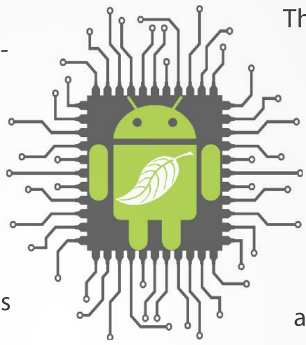


Conservation cores (c-cores): Small boxes in the image are the pattern of logic gates that are spatially placed over a small portion of the GreenDroid chip.

Taylor told *MIT Technology Review*, operating shrinking transistors with lower voltages was "traditionally the escape valve that enabled more computational power without more heat output, but now there is no place to go."

The payoff in energy efficiency (hence the "Green" in GreenDroid) is made possible because the chip design is based on information about a device's operating system and the most-accessed mobile apps. The GreenDroid prototype replaces some of the dark silicon with 100 or more small conservation cores, or 'c-cores,' which surround the central core of the processor. By tailoring each core to handle a specific app's code, c-cores can reduce by more than 90 percent the processor's energy use for the Android code that is targeted – making it the most efficient way of handling that one task.

"We believe," says Taylor, "that incorporating many automatically generated, specialized cores for the express purpose of saving energy is the next evolution in application processors."



The automated generation is built into the software developed by Taylor and his colleagues for taking the computational demands of popular apps and building a GreenDroid chip design with the new architecture.

So far, all of the GreenDroid results involve simulations, but the next step is to do benchmarking with a custom-fabricated processor, including features as small as 28 nanometers.

The first GreenDroid will be fabricated and ready for testing in 2013, with commercialization expected later in the decade.

* **GreenDroid: Exploring the Next Evolution in Smartphone Application Processors**, Steven Swanson and Michael Bedford Taylor, *IEEE Communications,* Volume 49, Number 4, pp. 112-119, April 2011. http://bit.ly/XHFa3L

C.K. Cheng's research is concerned with circuit analysis and physical planning of VLSI designs. For physical planning, his team constructs performance-driven layout systems to automate the interconnect planning, power network synthesis, data path generations, and packaging.

Rajesh Gupta works in microelectronic embedded systems on projects ranging from combating variability in semiconductors, non-volatile storage systems to energy efficiency in data centers and buildings. His past contributions include SystemC modeling and SPARK parallelizing high-level synthesis.

Andrew Kahng's research focuses on the VLSI design-manufacturing interface. He pioneered methods that link designer intent with the manufacturing process to improve yield and reduce power. He co-chairs the Design and System Drivers technology working group in the International Technology Roadmap for Semiconductors.

Ryan Kastner performs research in embedded computing systems, including reconfigurable computing, hardware security, underwater networks, robotics and biomedical imaging. He directs the UCSD-National Geographic Engineers for Exploration Program.

Pradeep K. Khosla's research areas include cybersecurity, Internet-enabled collaborative design, agent-based architectures for distributed design and embedded control, software composition and reconfigurable software for real-time embedded systems, distributed information systems, and more. He is also Chancellor of UC San Diego.

Embedded Systems Under the Sea

To create technology that aids in the collection of sensor data, CSE Prof. Ryan Kastner and his research group work closely with marine scientists and limnologists. The ultimate goal: a better understanding of underwater ecosystems, which in turn can shed light on issues ranging from global climate change to the impact of human behavior on coastal waters.

A major impediment to larger, higher-density underwater sensor networks is a cheap, low-power, wireless modem for underwater communication. Current underwater wireless modems cost thousands of dollars and are primarily used for deep-water, long-range communications. With funding from NSF, Kastner's team developed a low-cost (< $500), low-energy acoustic modem*. Designed from the ground up, each part – the transducer, analog and digital components, and software – was built to minimize the total cost and expenditure of energy. The team successfully tested two prototypes: one for moored applications that require a modem optimized for short ranges and low data rates; and a second prototype for a field that demands reliable communication in noisy, shallow-water environments, e.g., for the study of coral reef environments.

Kastner has also tackled new ways to detect objects underwater more accurately, with funding from the National Oceanographic and Atmospheric Administration (NOAA) and SPAWAR. Underwater environments are highly dynamic, with currents, surface weather and human involvement causing visibility to shift. In low-visibility situations, scientists typically employ a sonar-based system for vision, since sonar works regardless of visibility. Yet sonar has its own limitations when it comes to detecting or classifying an object.

When underwater conditions are good for visibility, video capture permits very robust computer-vision techniques for object detection and classification. The image analysis can classify based on pixel color or shape because of more defined edges. Unfortunately, when visibility is low – which is

Replacing the Flash in Memory

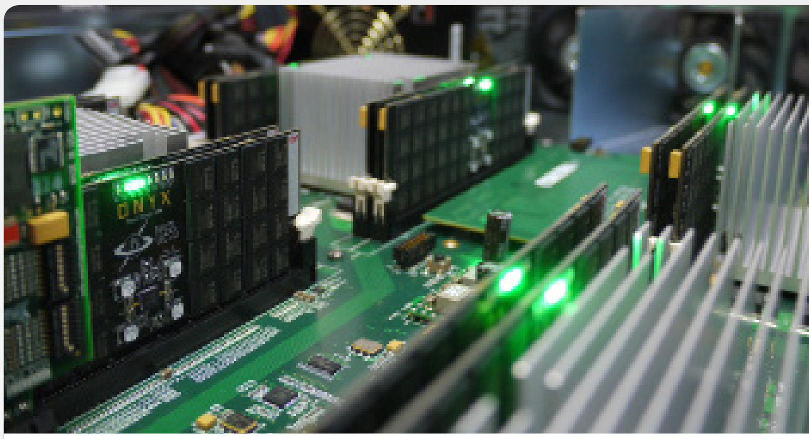
Today's computers are awash in data. Making sense of that data requires computer systems that can keep up, and one solution is to develop solid-state storage devices that give computers fast, energy-efficient access to the data they need.

Enter CSE's Non-Volatile Systems Laboratory (NVSL), which builds prototype storage systems that integrate cutting-edge memory technologies such as phase-change memory (PCM) and spin-torque transfer memories (STTMs). These prototypes include novel features that work hand-in-hand with specially engineered software components to reduce or eliminate performance bottlenecks left over from the disk-based storage systems of the past. The result is flexible computer systems that can help sift through vast amounts of data quickly and efficiently to support data-intensive applications such as Web search, social networks, and online fraud detection.

PCM and STTM aim to replace the flash memory that currently powers iPads, iPods and cell phones. They are faster and have a simpler interface than flash. They are more reliable as well, and they pave the way for data-processing systems powerful enough to tackle challenging problems from genome assembly and modeling the earth's climate to analyzing Web content in real time.

"We have found that replacing flash memory and disks will require more than simply swapping one memory technology for another," explains CSE Prof. Steven Swanson, who directs the NVSL. "When PCM and STTM enter the mainstream they will be fast enough to reveal hidden inefficiencies and latent bottlenecks in existing computer systems – systems optimized for disk-based storage systems that were up to 50,000 times slower compared to these emerging memory technologies."

To identify and eliminate these inefficiencies, Swanson and his team built the world's first, publicly-demonstrated PCM storage array, called Onyx*. Next, they used performance data from Onyx (and a related system called Moneta) to redesign the Linux operating system to match the level of performance that Onyx can deliver. In many cases, says Swanson, NVSL had to throw out existing components and start from scratch. In the process, they redefined how Linux provides programs with access to their data and how the system protects data from inadvertent corruption. The result: a 20-times improvement in performance for Moneta and Onyx, and a huge boost in performance at the application level.

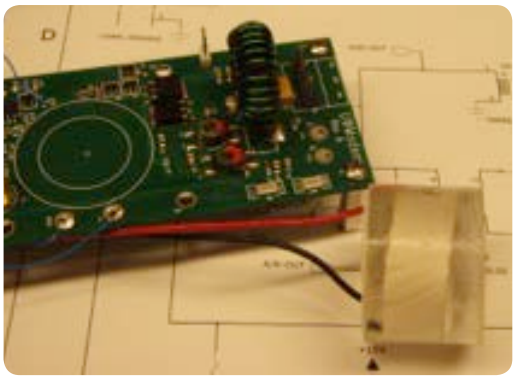

Interior of the Moneta storage array with custom-built Onyx phase change memory modules. It is up to seven times faster than commercial, state-of-the-art solid state drives, and thousands of times faster than conventional hark drives.

"Our version of Linux is ready for these new storage devices, but applications are not," says Swanson. "Databases also suffer from disk-related hidden inefficiencies and latent bottlenecks. "

The next step for Moneta and Onyx is to integrate key aspects of a database's functionality into hardware, streamline the software, reduce complexity, and improve performance."

For the last four decades, painfully slow disk-based storage systems have imposed a speed limit on the ability to process massive data sets. Moneta and Onyx are paving the way for new storage applications that will only be possible when storage is thousands of times faster than disk.

* **Onyx: A Prototype Phase-Change Memory Storage Array,** Ameen Akel, Adrian M. Caulfield, Todor I. Mollov, Rajesh K. Gupta, and Steven Swanson, *Proceedings of the 3rd USENIX conference on Hot Topics in Storage and File Systems*, June 2011 http://bit.ly/S5Oy0l


Low-cost underwater modem prototype

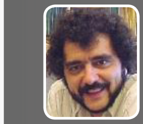
Scythe Butterfly fish detected using Kastner's hardware-accelerated system for detecting objects

more often the case – the video is not helpful until the vehicle is very close to the object in question.

Kastner and his colleagues believe that the future of underwater navigation based on vision must be an amalgamation of sonar and video data analysis. In low visibility, the sonar data can determine areas of interest, which can be further investigated by navigating closer to the object and utilizing computer-vision techniques. As visibility improves, the video data can be weighted more heavily in the algorithm.

* **Designing an Adaptive Acoustic Modem for Underwater Sensor Networks,** Lingjuan Wu, Jennifer Trezzo, Diba Mirza, Paul Roberts, Jules Jaffe, Yangyuan Wang and Ryan Kastner, *IEEE Embedded Systems Letters,* Volume 3, Issue 3, December 2011 http://bit.ly/YAn84k

Jason Mars works in runtime systems, datacenters, compilers and computer architecture, as well as online adaptive systems in both software and hardware. He earned his Ph.D. from the University of Virginia in 2012 with a thesis on "Rethinking the Architecture of Warehouse-Scale Computers."


Alex Orailoglu leads the Architecture, Reliability, Test group. His work in computer architecture spans embedded processors, application-specific processors, nanoarchitectures & reliable multicores. His contributions to VLSI Test include test compression methodologies, mixed-signal test, BIST diagnosis and adaptive test.

Tajana Simunic Rosing works at the intersection of software and hardware design. Her current research interests are in the area of low-power design of embedded wireless systems, thermal and power management of chip-multiprocessors and energy-efficient data centers.

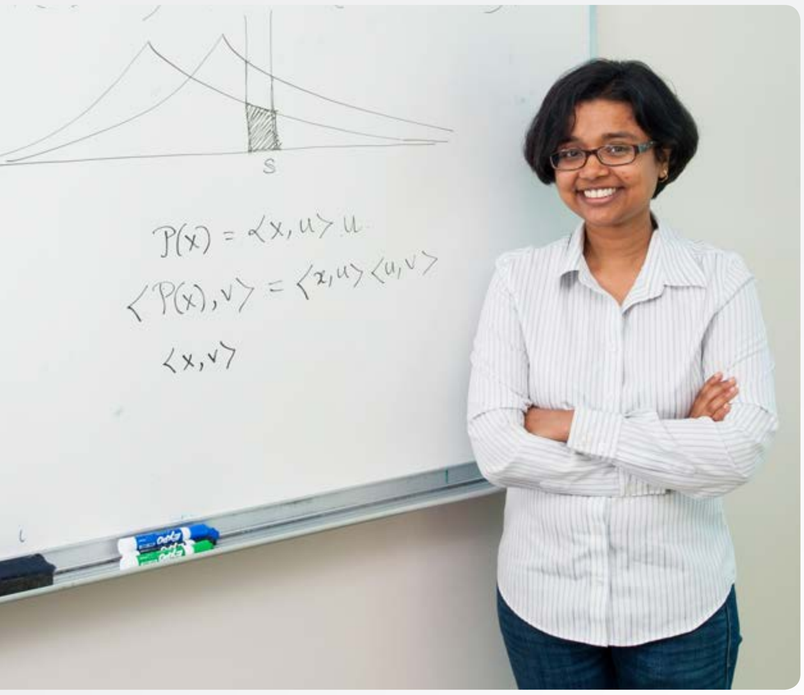
Steven Swanson is the director of the Non-volatile Systems Laboratory. His research interests include the systems, architecture, security, and reliability issues surrounding non-volatile, solid-state memories. He also co-leads projects to develop low-power co-processors for irregular, mobile applications (e.g., Android Apps).

Michael Bedford Taylor designs and builds novel hardware and software systems. Recent systems include: GreenDroid (facing page); Raw, a 16-core tiled many-core processor; and Kremlin, a tool that tells programmers which regions of their serial programs to parallelize.

Dean Tullsen does computer architecture and compilation and is best known for simultaneous multithreading in processors from Intel, IBM, Sun, etc. He helped introduce helper-thread prefetching, heterogeneous multicores, and recently focuses on enabling parallel speedup on serial code. He is an ACM and IEEE Fellow.

Quantifying the Price of Privacy

The data avalanche brought about by the digital revolution has made it possible to harness vast datasets for everything from statistical analysis to teaching machines to recognize patterns and respond in 'intelligent' ways.



But much of this data comes from humans, and many of those humans expect their data to remain private. Preserving this privacy, however, is not always easy, says CSE Prof. Kamalika Chaudhuri.

"Suppose you have a bunch of sensitive data, such as genomic data that you've gathered from patients, and now you want to compute some statistics on that data to develop some kind of prediction algorithm," she explains. "For example, you could be analyzing certain features of patients in order to predict if they might develop a certain disease."

"With most data-based research, so long as the patients' names and addresses are removed, the data is considered private," adds Chaudhuri. "But with datasets based on small sample sizes, this is not the case."

Chaudhuri and her colleagues have discovered that it is possible to 'reverse-engineer' the statistics obtained from such data to determine who the patients are, thus compromising their privacy.

To account for this, the researchers have developed a series of privacy-preserving techniques – collectively known as 'Differentially Private Empirical Risk Minimization'* – to determine how to classify data and subsequently develop prediction algorithms, while simultaneously maintaining privacy. One crucial aspect of the approach is to add a certain degree of noise to a data set to mask the effects of one person's data being added (so-called 'objective perturbation').

"But because you're adding a little bit of noise to the data, you're also losing some accuracy," notes Chaudhuri, "so we also try to quantify how much accuracy you're going to lose in order to provide privacy. The more samples in your data, the less the relative loss of accuracy, so accuracy is a function of sample size. So really what we're doing is quantifying the price of privacy."

Chaudhuri says her team's results show that, both theoretically and empirically, objective perturbation is superior to previous state-of-the-art techniques for managing the inherent tradeoff between privacy and learning performance. Better yet, the techniques can be used for any type of data – from medical data to financial data – thereby ensuring that machines can get smarter without compromising the human desire to remain anonymous.

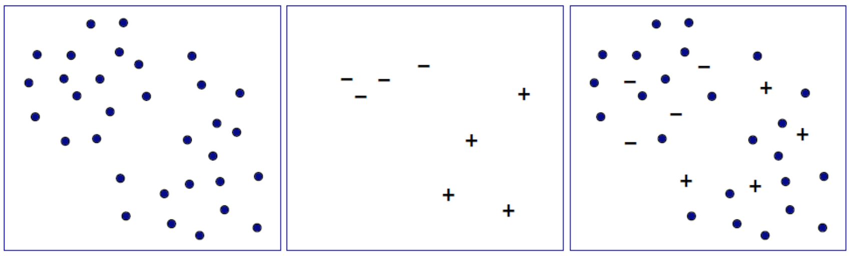
* **Convergence Rates for Differentially Private Statistical Estimation**, Kamalika Chaudhuri and Daniel Hsu, *Proc. 29th International Conference on Machine Learning*, June 2012. http://bit.ly/N3UBnm

Giving Computers a Confidence Boost

Sometimes it might seem like computers need humans like fish need bicycles. But the truth is, most computers still need humans – especially when it comes to teaching them the difference between fish and bicycles.

That's where CSE Prof. Sanjoy Dasgupta's research comes in. Dasgupta and his colleagues are developing techniques to enhance machine learning so that interactions between computers and humans are more efficient, and data mining and classification more accurate.

"There's a lot of data out there, so much so that there is no way that humans can process, label and interpret all of it," explains Dasgupta. "Computers can be used to sort through this data, but they do need human interaction to get a sense of what the data mean. That required human interaction is a limiting factor and ultimately causes bottlenecks, because human time is costly and there aren't that many of us, especially when it comes to areas that require some expertise."



Dots at far left represent 'unlabeled' points. Plus and minus signs (center) indicate labeling of only a few points using supervised learning, while ignoring the rest. At right, semisupervised and active learning get more use out of the unlabeled pool by using them to constrain the choice of a classifier, or by choosing informative points to label.

A computer, for example, can download millions of images from the Web, but it still requires some guidance from humans to be able to determine what is actually depicted. Teaching a computer to know what it's "looking at," says Dasgupta, is a necessary precursor for the computer to then recognize underlying patterns in subsequent datasets and be able to classify data without human intervention.

He is quick to point out, however, that most computers don't distinguish between digital images the way humans do. Humans are typically alert to distinguishing features, like the pointed ears, long tail and whiskers that signify 'cat.' A computer starts with pixels, and is interested only in converting that pixelated information into outputs that, based on previous examples it has learned, trigger its memory for 'cat' or 'not cat'.

Even a well-trained computer, however, can run into examples that give it trouble. Dasgupta's technique, known as 'active learning,' teaches computers to use their 'confidence' in determining what human feedback they need to distinguish between an image of a cat and a dog – or, say, Arnold Schwarzenegger as The Terminator, Conan the Barbarian or the former Governor of California.

Kamalika Chaudhuri's primary research area is learning theory, a broad field with connections to algorithms, statistics, artificial intelligence, information theory and signal processing. She is also interested in issues of data privacy.

Garrison W. Cottrell works on computational models of cognitive processes. He directs both the Interdisciplinary Ph.D. Program in Cognitive Science and an NSF-funded Temporal Dynamics of Learning Center, which has 40 PIs from 16 institutions.

Sanjoy Dasgupta develops algorithms for the statistical analysis of high-dimensional data. This is a broad, emerging research area with foundations in AI, information theory, probability/statistics, convex geometry, and theoretical computer science.

Applying Machine Learning to Computer Security

When you think about computer security, machine learning is not the first approach that springs to mind. Yet machine learning is the cornerstone of a novel strategy for detecting which URLs point to malicious Web sites — and ultimately, for protecting consumers from navigating to the online front of a criminal enterprise.

CSE Prof. Lawrence Saul traces the work on learning how to detect malicious URLs to spring 2008, when a Ph.D. student in computer security, Justin Ma, served as the teaching assistant for his undergraduate course in artificial intelligence. Saul joined fellow CSE faculty members Stefan Savage and Geoffrey M. Voelker as Ma's Ph.D. advisors. Ma – who is now at Google – earned his Ph.D. in 2010, and the following year he and his advisors published a complete description of their system for detecting malicious URLs in *ACM Transactions on Intelligent Systems and Technology**. The main result of this work was an online classifier that detects malicious Web sites with 99 percent accuracy over a balanced data set.

Lawrence Saul with small selection from among the millions of malicious URLs that lead consumers on the Internet to the online front of a criminal enterprise

The article explored how to detect malicious Web sites from the lexical and host-based features of their URLs. The researchers had access to a trove of malicious URLs made available to them by a large Web mail provider. An equally large number of benign URLs were obtained from certain online directories that had been previously vetted for validity.

In a nutshell, the machine-learning approach is to take a long list of URLs and extract features by asking simple questions about the tokens it contains (e.g., does it end in *.edu*, in which case it is less likely to be malicious than a *.com* address), and also how it was registered. There are potentially millions of questions to ask about the textual and host-based properties of URLs – and the answers to these questions change over time as malicious actors on the Web adapt to new defenses.

"You can look up many, many properties of a URL," explains Saul. "Not only how it is spelled, but also who registered it, the geographical location of the host, the server's connection speed, and so on. We're able to extract a million features from a single URL, and the job of our learning algorithm is to estimate a model that takes all these features as input and computes, for each URL, a single number that rates its maliciousness."

The research had enormous impact because the underlying model was capable of rapidly adapting itself to newly processed URLs. The researchers showed that the problem lends itself naturally to modern algorithms for online learning. Online algorithms not only process large numbers of URLs more efficiently than batch algorithms, which are limited by the number of training examples that can fit into memory, but they also adapt more quickly to new features in the continuously evolving distribution of malicious URLs.

"Spammers and other malicious actors are getting shut down every day, and there are always new ones popping up, so any static model becomes stale very quickly," says Saul. "Just consider the volume – billions of URLs per day – handled by an industry-scale system. We observed that a rapidly adaptive model makes a significant difference not just over days and weeks of operation, but even over hours."

The low error rate in detecting malicious URLs is largely attributed to the continuous retraining of classifiers in the face of new strategies developed by spammers and other online criminals. After publication of its 99 percent success rate, the algorithm was integrated into the infrastructure of a major Web mail provider.

The April 2011 paper looked only at URLs, but since then Saul and his colleagues have used the same machine-learning approach to vet text, images, and other content from malicious Web sites. Initial findings were included in a presentation on "Beyond Blacklists: Learning to Detect Malicious Web Sites from Suspicious URLs" during the UCSD Center for Networked Systems' spring 2012 Research Review.

Ultimately, the researchers would like to create a URL reputation service that leverages the power of machine learning on all aspects of the problem. According to Saul, the team also expects that its work will provide valuable lessons for other applications of machine learning to computer security.

* **Learning to Detect Malicious URLs,** Justin Ma, Lawrence Saul, Stefan Savage, and Geoffrey M. Voelker, *ACM Transactions on Intelligent Systems and Technology*, April 2011. http://bit.ly/Q8YRkn

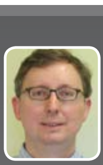
"The notion of 'confidence' is becoming a more important thing in computer systems," says Dasgupta. "For the iPhone to be a reliable and pleasant interface, Siri needs to be confident that it has understood your desires. This is something that is an open research area – developing systems that are not just willing to make guesses, but are also able to attach confidence values, i.e., 'I think it's a dog and I'm 90 percent sure.'"

Active learning* builds upon this measure of confidence, notes Dasgupta, and ultimately necessitates less human interaction than previous techniques.

"In active learning," he adds, "the computer starts out with an enormous amount of raw data, and then the human teaches it with just a few labels. Typically those are going to be the most informative labels, such as labels that point out something anomalous, or labels the computer is the least sure about. "

"When enough data is labeled," continues Dasgupta, "the computer can then determine what matches an established pattern and what doesn't. It learns a rule with an accompanying measure of confidence without having to ask for unnecessary help from a human."

* **Two Faces of Active Learning**, Sanjoy Dasgupta, *Theoretical Computer Science*, April 2011. http://bit.ly/FPSrof



Charles Elkan's research focuses on algorithms for learning from data, with applications in biomedicine, social science, and other fields. His earlier work on the foundations of fuzzy logic was the subject of Claude Rosental's book, "Weaving Self-Evidence: A Sociology of Logic."

Lawrence Saul works in machine learning and pattern recognition. He is best known for his work in high-dimensional data analysis and probabilistic inference. Saul is currently editor-in-chief of the *Journal of Machine Learning Research.*

Yoav Freund works on machine learning and its applications. His main focus is on confidence-rated prediction and on applications of machine learning in signal processing and control. Freund and Robert Schapire received the 2003 Gödel Prize and the 2004 Kanellakis Theory and Practice Award.

Protecting Anonymity in Location-Based Services

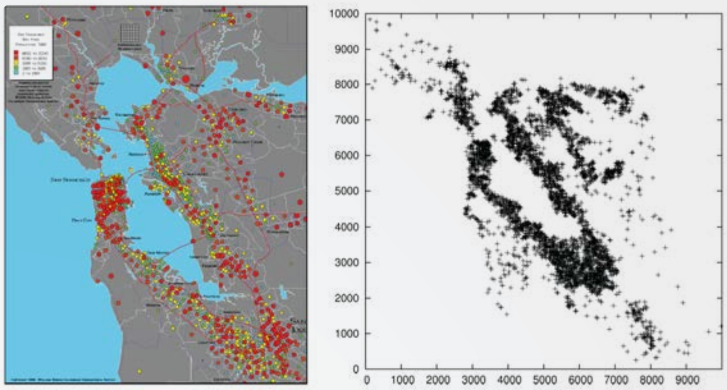
In a world where public records can be searched and downloaded with the touch of a keystroke, the cost of physically accessing, recording and tabulating this information has been a *de facto* barrier to its use and dissemination. The ease of access to everything from grocery-buying habits to patterns of political donations, and the development of powerful tools for the management and analysis of this data, have led computer scientists to explore ways to use such data while maintaining the anonymity and privacy of users.

Data holders typically strip their released data of identifying information, such as a street address or social security number. Yet CSE Prof. Alin Deutsch and other researchers believe that it is not unduly difficult to cross-reference pieces of 'unscrubbed' information with publicly available records to reduce the uncertainty about an individual's identity. More pieces of information could be suppressed or swapped within the data set, but such methods would also dilute the accuracy and usefulness of the data.

In order to protect privacy while also protecting the truthfulness of the data, software researchers developed the concept of 'k-anonymity' to hide the identity of mobile devices. K-anonymity dictates that released sets of data be such that any combination of values of information that could be used to identify people are indistinctly matched to at least 'k' respondents. This is accomplished through the generalization and suppression of certain pieces of data (e.g., a zip code can be generalized by eliminating the least relevant digit).

With the proliferation of wireless, location-based services (LBS), devices can be pinpointed so that users can access useful information relevant to their current location, e.g., finding highly-rated restaurants nearby. Wireless providers keep logs of queries that describe not only where a consumer is located, but also his or her interests and patterns of behavior – which can, in turn, be used to pitch products or services that fit a consumer's particular behavioral profile.

The CSE research team has shown that current methods for providing anonymity to LBS users are insufficient. According to Deutsch, the widely-used k-anonymity technique defends "only against naïve attackers who have no knowledge of the anonymization policy that is in use." In other words, when attackers gain access to a few key pieces of information – notably the request from the LBS log, the location of the mobile device, and the design of the system used to provide protection – the identity of the individual can be deduced.



Anonymizing for one million users in the San Francisco Bay area could be done with a delay of less than half a second by using just 16 servers.

An anonymization policy is a realistic threat, says Ph.D student Kevin Keliang Zhao, because "an attacker with subpoena powers, for example, a federal agency, or a disgruntled ex-employee, can obtain the design of the system." System designs are also based on well-accepted principles that are not secret.

According to Deutsch, the team's work* is "to keep the requestor's interests private even from attackers who, via hacking or subpoena, gain access to the request and to the locations of the mobile user and nearby users."

A top priority for a location-based service is speed, so any noticeable impact on responsiveness would likely meet with the disapproval of users. Because of the seriousness of the concern, Deutsch's team studied performance effects, and also concluded that the proposed algorithm results in only a minimum reduction in utility to users. The CSE researchers found that only 16 servers would be sufficient to provide anonymization for one million users in the San Francisco Bay area with a delay of under half a second – at only a 1% divergence in cost from the optimum. Deutsch concludes that their work "strikes a paradigmatic balance in the trade-off between strength of the privacy guarantee, utility, and running time for enforcement."

* **Policy Aware Sender Anonymity in Location Based Services**, Alin Deutsch, Richard Hull, Avinash Vyas, Kevin Keliang Zhao, *CoRR abs/1202.6677*, February 2012. http://bit.ly/X3DcPw

Scott Baden's research is in large-scale and parallel computation, and focuses on irregular problems and software techniques to improve programmer productivity. His work is interdisciplinary, and has resulted in new computational tools used in applied fields such as neuroscience, computational fluid dynamics and computational chemistry.

Alin Deutsch develops tools that assist database owners in publishing data on the Web, and developers in querying the integrated collection of such published data, conceptually treating the entire Web as a single, virtual database.

Jeanne Ferrante's research has been in high-performance computing, including scheduling in large-scale distributed systems. She is best known for her co-development of intermediate compiler representations, most notably Static Single Assignment (SSA) form. Ferrante is an ACM Fellow.

William G. Griswold works in ubiquitous computing, software engineering (especially refactoring), and human-computer interaction. His CitiSense project is developing technologies in support of participatory sensing of air quality for personal health.

Ranjit Jhala is interested in programming languages and software engineering, and more specifically in techniques for building reliable computer systems. His work draws from, combines and contributes to the areas of model checking, program analysis, type systems and automated deduction.

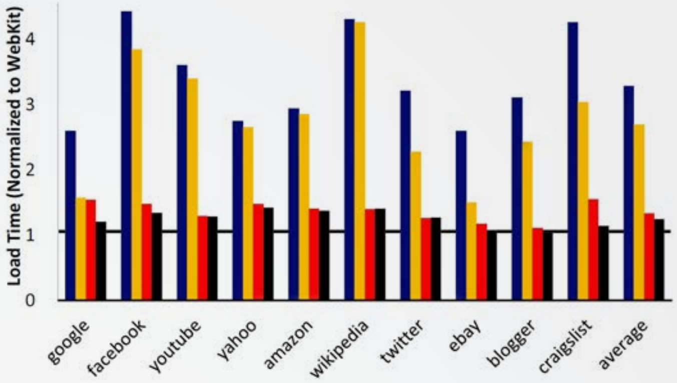
Scott Klemmer joined CSE in 2012. Organizations around the world use Klemmer's open-source design tools and curricula. He has written on the value of continuous feedback to enhance the quality of crowd-sourced expertise.

Does Quantified Health Compute?

To the *Atlantic Monthly*, he is "The Measured Man." To the *San Diego Union-Tribune*, Larry Smarr is a standard-bearer for the emerging 'Quantified Self' movement. And as a reporter writing in *MIT Technology Review* put it, Smarr "has become a poster man for the medical strategy of the future." The CSE professor is the founding director of the California Institute for Telecommunications and Information Technology (Calit2). In his own research, he works at the intersection of high-performance, distributed computing and the proliferation of devices, tests and DNA sequencing that allow Smarr to track his own health in minute detail. Calit2 has allowed him to envision the future of body imaging in its StarCAVE virtual-reality environment, where he takes visiting physicians on a 3D fly-through of his own gut. Or he runs through a battery of physical tests in the Exercise and Physical Activity Resource Center (pictured at right), and pores over massive data sets from the sensors he wears 24/7. "The data available from MRIs, EEGs and other diagnostic tests and imaging is over one million times what it was just a decade ago, and it's growing exponentially," explains Smarr. "How we turn all of that data into proactive health decisions or medical diagnoses is as much a challenge for computer scientists as it is for medical professionals."

Designing a Secure Browser

The popularity of the Web has in turn made the browser into the principal software platform for interacting with Internet content. For the same reason, browsers have also become the principal target for attackers. Unfortunately, current browsers are fragile They are complex pieces of software with rich features that allow for flexibility and programmability, and even small bugs can make the browser vulnerable to attack. CSE Prof. Sorin Lerner and co-authors Dongseok Jang and Zachary Tatlock explore a new approach to secure browser design in a paper delivered at the 21st USENIX Security Symposium in August 2012*. They explain that previous verification techniques for browser security operate on a *model* or *abstraction* of the browser, and not on its actual implementation. This has created what Tatlock and Jang call a 'formality gap,' a discrepancy between what is verified and what is implemented. It is through this gap that hackers can infiltrate a browser even if it has been verified using strong formal methods.



Quark load times for the Alexa Top 10 Web sites, normalized to stock WebKit's load times. In each group the blue bar shows the unoptimized load time, the black bar shows load time in the final, optimized version of Quark, and center bars show how additional optimizations improve performance.

There is one known way to bridge this formality gap: implement the software in a proof assistant and use the proof assistant's interactive environment to formally prove, in full formal detail, that the software implementation is correct. However, building formal proofs in full detail for realistic applications with millions of lines of code is extremely time-consuming, if not completely impossible.

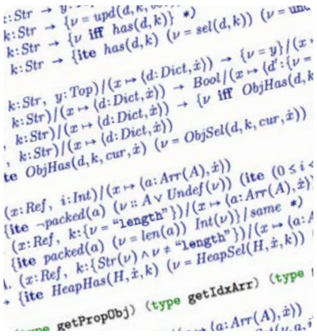
Lerner and his colleagues devised a technique, dubbed 'formal shim verification,' which restricts the number of lines of code that must be verified to a few hundred rather than a few million. Formal shim verification, says Lerner, consists of "creating a small browser kernel which mediates access to security sensitive resources, and then formally verifying that this browser kernel is correct using a proof assistant."

To demonstrate the idea, the team created Quark, a Web browser that uses a kernel-based architecture similar to Google Chrome. Unlike Chrome's kernel, however, the Quark kernel has been formally verified using a proof assistant from basic axioms, thus providing strong guarantees of security, even if the rest of the browser uses state-of-the-art implementations that have not been verified. In this way, Quark is able to use the unverified WebKit layout engine – the same engine used in Safari and Chrome. Using such realistic components has made Quark into a practical, usable and secure browser, which can successfully run complex pages like Gmail, Google Maps, Facebook and Amazon.

Although Quark runs such complex pages, the browser is still in the prototyping phase. It does not yet support some standard features of the Web, such as third-party cookies, and in some cases it enforces non-standard security policies. Tatlock, Jang and Lerner are determined that this is only the first implementation of Quark. They already have some ideas about how to include a number of standard browser features without severely complicating their kernel – or having to work through a fundamental redesign.

* **Establishing Browser Security Guarantees through Formal Shim Verification**, Dongseok Jang, Zachary Tatlock and Sorin Lerner, *Proc. 21st USENIX Security Symposium*, August 2012. http://bit.ly/XMs5ZS

Making JavaScript Web Applications More Efficient


```
::Str → y...
k:Str → {ν = upd(d,k, ...
k:Str → {ν iff has(d,k)} *)
k:Str → {ite has(d,k) (ν = sel(d,k)) (ν = un...
k:Str, y:Top)/(x ↦ ⟨d:Dict,ẋ⟩) → {ν = y}/(x ...
k:Str)/(x ↦ ⟨d:Dict,ẋ⟩) → Bool/(x ↦ ⟨d':{ν = ...
k:Str)/(x ↦ ⟨d:Dict,ẋ⟩) → {ν iff ObjHas(d,k
, k:Str)/(x ↦ ⟨d:Dict,ẋ⟩)
te ObjHas(d,k,cur,ẋ) (ν = ObjSel(d,k,cur,ẋ))
(x:Ref, i:Int)/(x ↦ ⟨a:Arr(A),ẋ⟩)
{ite ¬packed(a) (ν :: A ∨ Undef(ν)) (ite (0 ≤ i <
(x:Ref, k:{ν = "length"})/(x ↦ ⟨a:Arr(A),ẋ⟩)
{ite packed(a) (ν = len(a)) Int(ν)}/same *)
(x:Ref, k:{Str(ν) ∧ ν ≠ "length"})/(x ↦ ⟨a:A
{ite HeapHas(H,ẋ,k) (ν = HeapSel(H,ẋ,k))
getPropObj) (type getIdxArr) (type
⟨a:Arr(A),ẋ⟩)
```

Due to its many compelling features, the JavaScript (JS) language has transformed the way in which software systems are developed, deployed and extended. JavaScript's 'dynamic' features allow for a great deal of flexibility and code re-use, permitting developers to lash together different components quickly. The language is also very 'mobile.' "HTML tags and JavaScript's eval facility can be used to download and execute code from the network, making it easy to enrich applications by flexibly incorporating library components hosted on diverse Web sites," explains CSE Prof. Ranjit Jhala. "Consequently, JS is now used to build complex, security-sensitive applications for communications, retail and banking, and is even a primary building block of Web browsers themselves."

Unfortunately, the advent of JS has also opened the door to new classes of security vulnerabilities, as applications manipulate security-critical client information such as browsing history, passwords, bank account numbers, social security numbers, and so on. Worse, the absence of any language-level isolation mechanisms makes it hard to ascertain the safety and security of components loaded from the Web, and to prevent these components from behaving in an undesirable manner.

Now a CSE team led by Jhala has presented a novel technique* for making JS programs more reliable, secure, maintainable and efficient, by developing a 'type' system that can accurately predict, at development time, the kinds of values that different JS functions and variables can take on at run-time. Previous attempts to design such a type system have fallen short as JavaScript's dynamic features, such as run-time 'reflection' and extensible dictionaries ensure that a variable's type changes as the program executes, thereby blunting the classical tools in the designer's kit.

The key insight of Jhala and his colleagues was that while the type of 'individual' variables rapidly changes as the program executes, the relationships between multiple variables remain quite stable and can be used to characterize the behavior of different parts of the program. "To formalize this insight," says Jhala, "we developed a dependent-type system which encodes the relationships using efficiently decidable logics, and verifies the program with a novel type-checking algorithm based on recent advances in logical constraint solving."

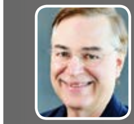
To evaluate their system, the CSE team built a checker and used it to precisely analyze a number of challenging programs from various sources, including the popular book "JavaScript: The Good Parts" and the SunSpider benchmark suite. Going forward, the team will build on their system as a foundation for developing automated and extensible security and reliability analyses for large JavaScript code bases.

* **Dependent Types for JavaScript**, Ravi Chugh, David Herman and Ranjit Jhala, *Proc. 27th Annual ACM Conference on Object-Oriented Programming, Systems, Languages, and Applications*, October 2012. http://bit.ly/PBuwrl

Ingolf Krueger and his team are developing methods, tools and infrastructures for service-oriented software and systems engineering, with applications in health, energy and cyber-physical systems. He directs the Gordon Engineering Leadership Center and co-founded LonoCloud, Inc.

Sorin Lerner is interested in programming languages and analysis techniques for making software systems easier to write, maintain and understand. He has worked on extensible and verified compilation frameworks, static and dynamic analyses for Javascript security, and scalable analyses for race detection.

Yannis Papakonstantinou works at the intersection of data management and the Web, including information integration, rapid development of cloud-based Web and mobile data-driven applications, and querying of semi-structured data. His work on enterprise information integration led to one of the first industrial XML-based virtual database platforms.



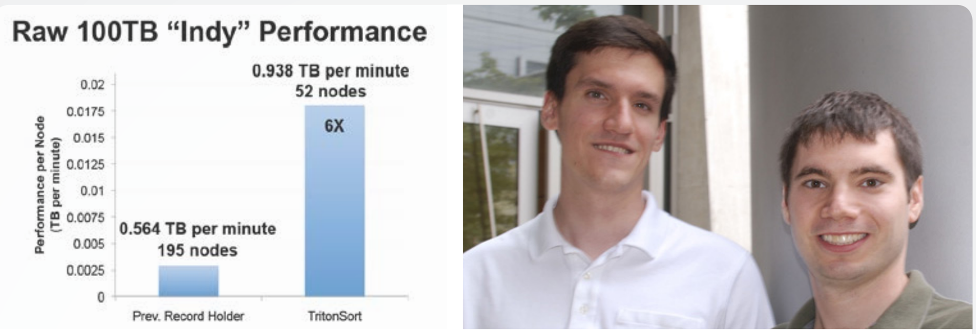
Larry Smarr is the founding director of the California Institute for Telecommunications and Information Technology (Calit2) and holds the Harry E. Gruber Chair in CSE. He has driven major developments in information infrastructure, including high-performance computing, the Internet, scientific visualization, virtual reality, global telepresence, and quantified health.

Victor Vianu works in database theory, logic and complexity, and data management on the Web. He is best known for his work on the theory of query languages and static analysis of XML queries and formal verification of data-centric Web services and business processes.

Off to the Races

TritonSort is a case study in building balanced, data-intensive and scalable computing systems. The system's staged, pipeline-oriented sorting system was designed to sort data with a very high degree of per-node efficiency. As a team of CSE researchers proved at the Sort Benchmark data-processing competition in 2012, TritonSort can realize orders-of-magnitude improvements in throughput and efficiency.

CSE researchers affiliated with the Center for Networked Systems have set seven world records in the annual competition for large-scale data processing, five of which remained world records after the 2012 contest.



(Left) Using TritonSort, CSE researchers sorted 100 Terabytes of raw data six times faster than the previous Indy record holder, using one-quarter of the nodes. (Right) Ph.D. students Rasmussen and Conley.

The Sort Benchmark competition is considered the Formula One and Daytona 500 rolled into one – but applied to fast data rather than fast cars. It attracts competitors from academic and industry labs all over the world, who vie to implement ever-faster datacenter designs.

The UC San Diego team includes CSE Prof. Amin Vahdat, Research Scientist George Porter, and Ph.D. students Alex Rasmussen and Michael Conley. Their TritonSort system* 'raced' in the computing equivalent of the high-end Indy 500, as well as the general-purpose Daytona category. In both competitions, TritonSort sorted one terabyte (1 trillion bytes) of data in as little as 106 minutes (Indy) and 138 minutes (Daytona).

TritonSort achieved record speeds by focusing on per-disk and per-node efficiency. Ultimately, its goal has been to sort data at the speed of the disks by keeping all disks constantly reading or writing data in large, contiguous chunks.

Designing a system to compete in the Indy category was comparable to constructing a racing vehicle that could only be driven on a track. After doing so well in that category in 2010, the CSE team decided to apply some of their improvements to a general-purpose version of TritonSort to 'race' in the Daytona category. They made improvements to the data structures and algorithms to make the system a lot more efficient in terms of sending records across the network. It did so by implementing MapReduce on top of its core components.

The key to the TritonSort design, says Porter, is seeking an efficient use of resources: "The whole aim of this project is to build balanced systems."

Although they easily won the top categories in 2012, the team fell short of beating its own world record set in 2011 in the Indy category.

Beyond speed, the efficiency of TritonSort's design drew raves. While the runner-up in 2011 used 3,500 nodes to achieve their second-place result, the UCSD team's entry used only 52 nodes. If implemented in a real-world datacenter, that means that a company using TritonSort could sort data more quickly, while only making one-seventh of the investment in equipment, space, and energy costs for cooling and operation.

The TritonSort team also came out on top in a relatively new competition, the 100-Terabyte Joulesort, in which teams vie to build a system that can sort the greatest number of data records while only consuming one joule of energy. (It takes a couch potato roughly one million joules just to watch TV for an hour.) The introduction of this category reflects a growing challenge facing large datacenters: energy efficiency. In the 2012 competition, TritonSort placed first in the 100-TB Joulesort in both the Daytona and Indy categories (in the latter category, sorting 9,700 records with just one joule of energy). A primary reason that datacenters are expensive to operate is because of the staggering scale of their energy consumption. Any design that can increase energy efficiency will have a positive impact on both the environment and on a company's bottom line.

Not content to rest on their laurels, TritonSort team members are now developing Themis, the platform on which TritonSort was built. It is being enhanced as part of the ongoing effort to apply the lessons learned when building TritonSort to more general-purpose, data-intensive, scalable computing workloads. Indeed, the TritonSort version that did so well in the Daytona categories in 2011 and 2012 was actually ThemisMR, an early version of the framework, with MapReduce built on top to enhance the system's balance, speed and robustness.

* **TritonSort: A Balanced Large-Scale Sorting System**, Alexander Rasmussen, George Porter, Michael Conley, Harsha Madhyastha, Radhika Niranjan Mysore, Alexander Pucher and Amin Vahdat, *Proc. 8th USENIX Symposium on Networked Systems Design and Implementation*, March 2011. http://bit.ly/10CZIzl

Datacenter Networks

The growth of complex, data-driven Internet applications such as Web search, social networking, interactive maps and streaming video has accelerated at an astounding rate. Today a single such application may harness hundreds of thousands of servers, all working together. Keeping pace with this new workload has required developing a new generation of communications networks designed specifically for the datacenter environment. These networks must deliver sufficient bandwidth to these resource-hungry applications, while ensuring that they are scalable, manageable, flexible, and have low-cost and low-power requirements.

This new domain of datacenter network design is one in which UC San Diego has taken a leadership role, designing and building new network switches that have been highly influential in industry and academia alike.

One of the key challenges in datacenter network design is that while it is possible to grow computing and storage incrementally by simply buying more servers and more disks, growing the network has required increasingly specialized and expensive switches. These traditional, 'scale-up' designs require that when more servers need to be supported by the network, expensive core switches must be replaced with yet faster switches. These faster switches are very complex, require significant amounts of energy, and can be prohibitively expensive.

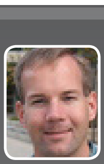
In 2008, CSE Prof. Amin Vahdat and his students used a network topology called a 'fat tree' to establish the practical 'scale-out' network design approach. Scale-out networking relies on providing multiple paths

Center for Networked Systems research scientist George Porter with the new Mordia switch located at Calit2

Keith Marzullo works on theoretical and practical issues of fault-tolerant distributed computing in various domains, including grid computing, mobile computing, and clusters. He also studies the relation between reliability and security.

Joseph Pasquale does research in operating systems, distributed systems and networks, focusing on performance and reliability of Internet-scale systems with highly decentralized control. He generally teaches operating systems as well as a math and engineering study-abroad program in Rome, Italy.

Stefan Savage works at the intersection of systems, networking and computer security. He is particularly fascinated by the interplay between the technical and socio-economic factors of problems, but is a rabid empiricist and likes measuring things. Savage is a Sloan Fellow and ACM Fellow.



Alex C. Snoeren's research interests include operating systems, distributed computing, and secure wireless and wide-area networking. His recent work focuses on emerging challenges in mobile and cloud computing. Snoeren received a Sloan Research Fellowship in 2009.

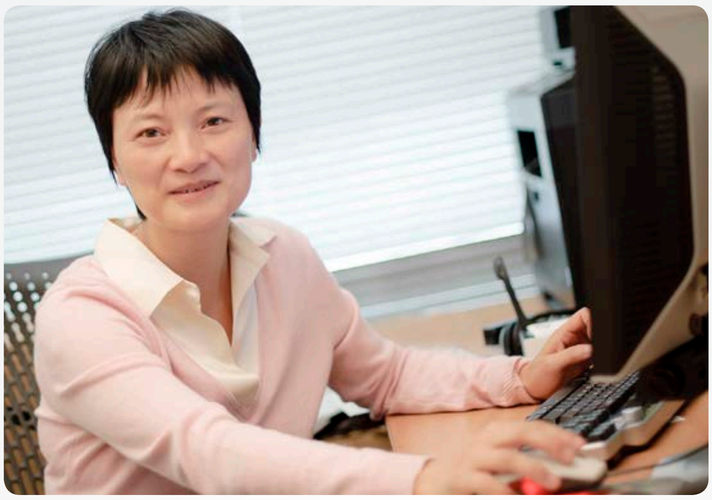
Taking the Cringe Factor Out of Computer Crashes

Computing systems fail. It's a fact of life. The trick, says CSE Prof. Yuanyuan (YY) Zhou, is to design around system failures so they are easy to diagnose and, ideally, able to be tolerated.

"When we talk about failures, we're talking about when a program crashes, or when the operating system goes to a blue screen, or when Google or Gmail is suddenly not available," explains Zhou. "Google's datacenter, for example, has a system failure every one-to-two seconds on average. The more machines a datacenter has, the more chance there is for things to fail."

It's a problem that affects just about every aspect of the tech industry, from storage system providers to data warehousing and network infrastructure companies.

"Storage failure is a really important issue, because a failure could mean that data is actually lost, not just that it isn't accessible," says Zhou, who co-authored three papers in 2012 on 'proactive' logging, including one delivered at the USENIX Symposium on Operating Systems Design and Implementation in October 2012*. "Their customers include a lot of banks, for example, so losing data is a big deal. And in the case of a network infrastructure company, it's the backbone of the Internet, so when there is a failure, information just isn't being communicated."



The problem with most software companies and service providers is that they design for functionality and performance and focus on what should happen when things go right. When something goes wrong, most companies rely on the end user to send an 'error report' in an attempt to diagnose the failure.

However, many users don't like sending error reports back, because it could compromise their privacy, and sending an error report won't necessarily help the end user with the problem they are currently experiencing. After all, most reports are used to diagnose errors so they can be eliminated from future versions of the software.

With funding from NSF as well as Intel, Cisco Systems, NetApp, Motorola, IBM and Teradata, Zhou and her colleagues have developed a suite of software tools to help companies detect defective software bugs and automatically – and remotely – log and self-diagnose errors through enhancements to code.

Developing the software meant enlisting the help of three major storage, data warehousing and network infrastructure companies, which provided access to their error data – something Zhou calls "a luxury."

"Industry has provided us a lot of support," she adds, "because they see the practicality of our tools."

Zhou and her team then used data mining and compiler techniques to detect discrepancies between intention (what a program is trying to do) and implementation (what actually happens). To improve methods for anticipating failure, they then determined which types of diagnostic information should be collected, and then automatically generated test cases – even test cases that included user error.

"In many cases, software is not tested against legitimate mistakes made by users who are not sophisticated," Zhou points out. "What we've done is harden the system against user mistakes and operator mistakes, so even if the user makes mistakes, the system can handle it gracefully."

Zhou says the next step for her and her colleagues is to release a set of similar tools for Android-based smartphones.

"Many smartphone apps are free for download but a lot of them also have defects, and that can eat a lot of battery power," she explains. "This problem is only going to become more and more prevalent because apps are cheap to make, but not everyone making them is testing them. Our tool will tell the user which app is draining their battery and how to fix it."

Their current suite of tools has already proven popular among users such as Qualcomm, Cisco, EMC, Juniper and Motorola, and not only because it spares them potential headaches. It also saves them money. Says Zhou: "Being able to detect, diagnose and log errors can save companies tens of millions of dollars a year."

* **Be Conservative: Enhancing Failure Diagnosis with Proactive Logging**, Ding Yuan, Soyeon Park, Peng Huang, Yang Liu, Michael M. Lee, Xiaoming Tang, Yuanyuan Zhou and Stefan Savage, *Proc. 10th ACM/USENIX Symposium on Operating Systems Design and Implementation*, October 2012. http://bit.ly/TXIF3S

through the network to add more bandwidth incrementally, rather than replacing them with faster paths. The result is that more servers can be supported by simply adding proportionately more low-cost, energy-efficient switches. Follow-on projects Portland (SIGCOMM '09) and Hedera (NSDI '10) extended this design to support existing Ethernet and IP protocols, enabling standard Internet applications to scale to hundreds of thousands of nodes, while remaining easy to manage and resilient to failures and outages.

Today the scale-out networking approach has been widely adopted in industry and it has driven the creation of an entirely new market segment for efficient datacenter networking. Yet the group has not stopped there. Since 2010, Vahdat and Center for Networked Systems research scientist George Porter have been exploring how datacenter networks might combine traditional, packet-based switching with the circuit-based switching offered by opto-electronics and fiber optics. Compared to all-electronic networks, such hybrid designs can support far higher speeds, while minimizing energy requirements and maintaining compatibility with existing network protocols.

Making such an architecture practical, however, requires correctly configuring and scheduling traffic from the electrical portion of the network to the optical. "The key challenge in marrying optics and electronics is minimizing switching time," says Porter. "Short-lived circuits allow handling more dynamic traffic, but they also create new challenges."

The first generation of their switch architecture, Helios (SIGCOMM 2010), was able to reconfigure new circuits in 15 milliseconds. While this provided huge scalability benefits for stable traffic patterns, it was not sufficiently quick for highly dynamic datacenter workloads. The next generation Mordia switch provides a separate wavelength to each rack and can completely reconfigure in under 12 microseconds* – three orders of magnitude faster – and can support a full range of modern applications.

* **Hunting Mice with Microsecond Circuit Switches**, Nathan Farrington, George Porter, Yeshaiahu Fainman, George Papen and Amin Vahdat, *ACM HotNets*, October 2012. http://bit.ly/S5CLjT

Amin Vahdat's research encompasses networks, distributed systems, and operating systems. He focuses on scale and reliability in datacenter networks and infrastructure applications. Vahdat has received an NSF CAREER award and a Sloan Fellowship.

Geoffrey M. Voelker enjoys working on fun and interesting problems in systems, networking, security, and wireless. His projects have explored Internet cybercrime, wide-area storage services, and local-area wireless networks.

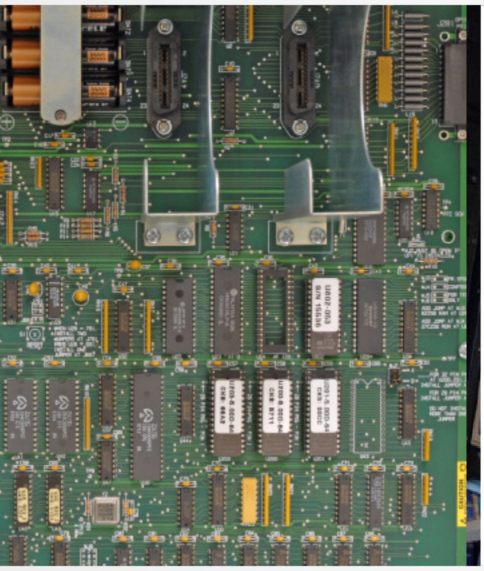
Yuanyuan (YY) Zhou's research interests include operating systems, networking, reliability and large-data analysis. Her research passion has focused on techniques for analyzing system data to improve software quality, manageability and reliability.

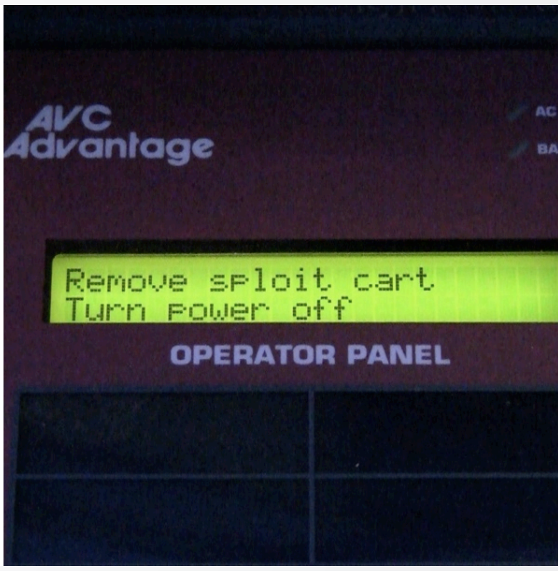
Hacking Democracy

As Joseph Stalin infamously said, "I consider it completely unimportant who in the party will vote, or how; but what is extraordinarily important is this — who will count the votes, and how." The Digital Age version of that statement revolves around the multiplicity of computerized systems used in capturing, reporting and tabulating votes. Indeed, as our democratic process has come to depend on a trusted tier of hardware and software, the security concerns of these systems have become ever more important. CSE Prof. Hovav Shacham has been one of the researchers leading this examination.

Shortly after joining UC San Diego in 2007, Shacham was part of the team commissioned by then-California Secretary of State Debra Bowen to complete a top-to-bottom review of the state's electronic voting systems. With colleagues from Stanford, Rice and the private sector, Shacham analyzed the source code for the widely-used Hart InterCivic system. He identified a range of vulnerabilities that could potentially allow votes to be manipulated, and the secrecy of a citizen's vote to be compromised – ultimately leading Secretary Bowen to decertify the machine in California elections.

Like several past studies, the findings were widely criticized by the voting-machine industry as being unrealistic (since the researchers had access to the underlying source code), and impractical, because they allowed arbitrary access to the machine itself. To address these claims and make concrete the threats of tampering, Shacham, his then-Ph.D. student Stephen Checkoway, and CSE staff member Brian Kantor, decided to investigate what could be accomplished *without* such capabilities. To do so, they targeted another popular voting machine – the AVC Advantage. This particular machine was designed to be impervious to most attacks because the voting software was stored in read-only memory (ROM) and the processor interface was specifically designed to prevent the execution of instructions except from this ROM. Thus, in this design, it is impossible to inject new code and cause it to be executed. However, two years earlier Shacham had pioneered a new attack strategy called return-oriented programming*, or ROP, that constructs malicious code entirely from snippets of benign programs, potentially bypassing such protections. Working with colleagues at Princeton University, Shacham and his team gained access to an AVC machine, reverse-engineered its ROM, developed an ROP-based attack, and embedded it in a standard voter cartridge – allowing complete control over election results with only 30 seconds of access to a machine.





Since then, Shacham has continued to examine the election ecosystem, collaborating with vision researchers to produce OpenScan**, the first ballot-counting system that allows competing parties to tabulate votes together and with statisticians to explore how to best validate contested elections.

* **Can DREs Provide Long-lasting Security? The Case of Return-Oriented Programming and the AVC Advantage**, S. Checkoway, A.J. Feldman, B. Kantor, J.A. Halderman, E.W. Felten and H. Shacham, *USENIX EVT/WOTE*, 2009. http://bit.ly/YgVAp7

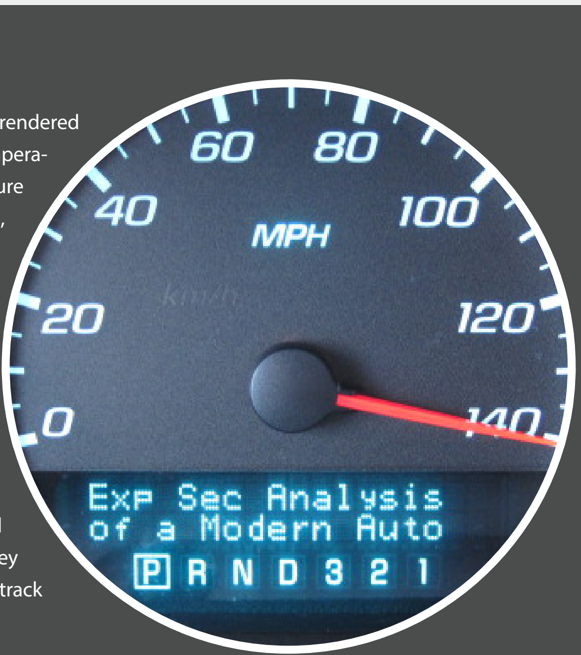
** **OpenScan: A Fully Transparent Optical Scan Voting System**, K. Wang, E. Rescorla, H. Shacham and S. Belongie, *USENIX EVT/WOTE*, August 2010. http://bit.ly/WqU1Qg

The Ghost in the Machine

While we primarily think of automobiles as analog mechanical devices, their control has long since been surrendered to digital computers. Indeed, a broad range of functions in the modern family sedan – from braking to temperature control – are managed by a network of 30 or more microprocessors working in concert. This architecture has without doubt increased safety (e.g., anti-lock brakes) and increased efficiency (precision fuel injection, for instance), but also may expose new risks that are only exacerbated by a parallel trend to add ever more external digital communication channels to vehicles (e.g., Bluetooth, telematics, digital radio, Wi-Fi, etc.).

It was precisely these transformations that led CSE Professors Stefan Savage and Hovav Shacham to investigate the security issues exposed in modern automobiles. Working with University of Washington professor Yoshi Kohno (himself an alumnus of the UCSD CSE Ph.D. program), they purchased late-model automobiles and set their students to understanding the role played by computers in such vehicles and how they might be compromised. In a pair of landmark papers*, the team showed that the Electronic Control Units (ECUs) in modern vehicles are highly susceptible to attack – allowing arbitrary control of safety-critical features such as lights and braking. The researchers also showed that attackers can mount and control such attacks remotely without any physical access to the vehicle. In a series of associated videos, they demonstrated the ability to remotely engage or disable the brakes, start and stop the car, disable door locks, track the car's location, and surreptitiously stream live audio from within the passenger compartment.

Amplified by coverage in both the *New York Times* and the National Academies' recent report on "The Safety Challenge and Promise of Automotive Electronics," the CSE research has already had major impacts in the automotive industry with the creation of new working groups on standards initiated by the Society of Automotive Engineers (SAE) and



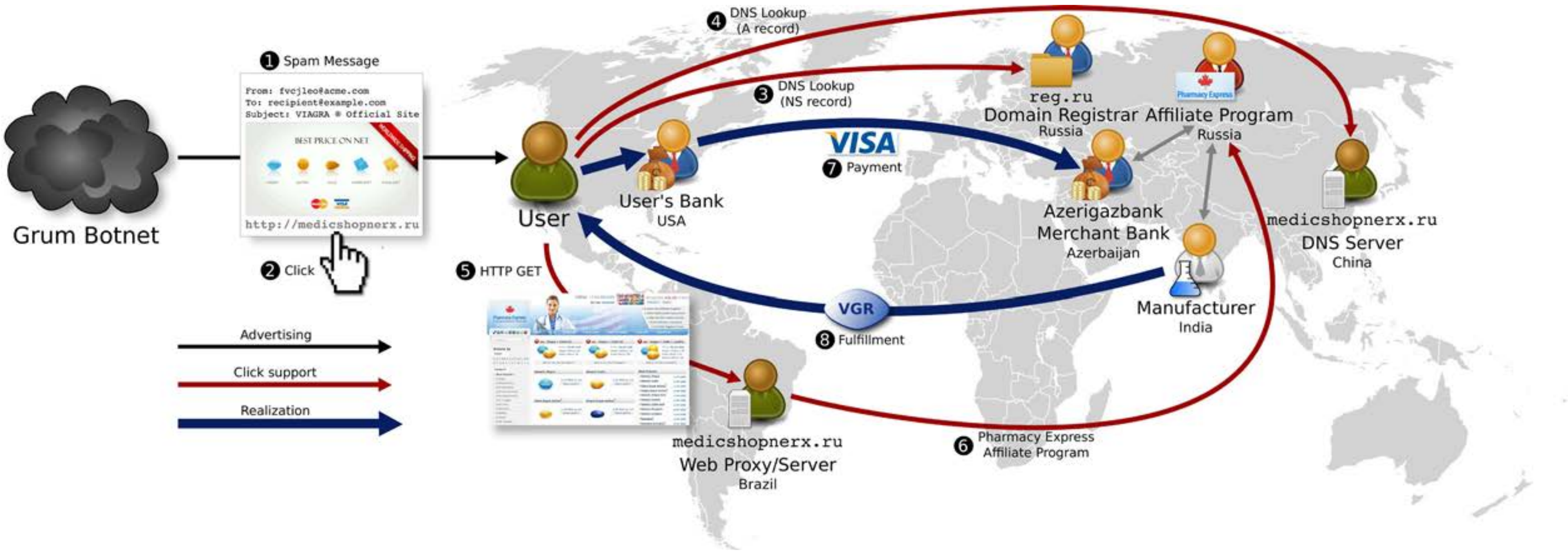
Displaying an arbitrary message and false speedometer reading (note that car is in park)

Understanding the Spam Value Chain

Traditionally, computer security has been viewed as a problem of technical origins. Our systems would be secure, "but for" deficiencies in their design or implementation. However, this viewpoint ignores the role of people in the security equation, both attackers and victims. "Trying to solve the computer security problem by fixing bits of software is like trying to solve the illegal drug problem by building a better border fence," says CSE Prof. Stefan Savage. "Sure, you need to understand how you're getting attacked, but every bit as important is 'why' and 'by whom'."

Savage is part of a group, including CSE Prof. Geoff Voelker, Research Scientist Kirill Levchenko, their Ph.D. students and long-time collaborators at the International Computer Science Institute at Berkeley, who together focus on the interplay between technical security issues and their economic and social underpinnings.

Most recently, they have turned their attention to examining abusive advertising strategies, such as e-mail spam. While billions of dollars are spent trying to block such messages, the CSE researchers instead looked at how spammers operate and how they monetize their investments. In pursuit of this goal, they tracked close to one billion spam e-mail messages, visiting each distinct site to identify its domain registrar, name servers and Web servers. They then used the content to cluster the sites together according to their sponsoring 'affiliate programs' – business entities that pay spammers on a commission basis. The group then placed hundreds of orders from these sites, leveraging a relationship with credit card issuers to identify the merchant banks being used to receive payment. "The goal of all this effort," says Voelker, "is to identify the full set of resources required to monetize spam-based advertising, and to learn which of these are the most valuable and hence most sensitive to disruption." (The figure above shows the set of resources used by just one such spam-advertised Web site.)



Their resulting "Click Trajectories" paper* at the 2011 IEEE Symposium on Security and Privacy demonstrated that, while resources such as domain registration, name servers and Web hosting are plentiful and cheap to replace, the underlying credit-card payment systems were not. Indeed, only three banks were used to monetize over 95 percent of all the spam e-mail messages they studied. "Ultimately, all of this activity is funded by consumer payments," says Levchenko, lead author of the study. "It is this money that in turn funds the 'cost centers' of the underground ecosystem, including botnets, malware, bullet-proof hosting, etc."

Based on these findings, a number of brand holders, working together with payment card networks, have tried to focus on this payment processing 'choke point' using the same techniques as in the study. In the group's most recent work, they tracked the efficacy of these efforts by pairing underground monitoring, hundreds of contemporary purchases, and information from brand holders and financial institutions about their activities. In their subsequent paper** at the 2012 ACM Conference on Computer and Communications Security, they documented the decimation of the online counterfeit software market and deep injuries to the market for online, unlicensed pharmaceuticals. Dozens of programs have closed and those remaining have lost significant profitability as it has become harder for them to accept credit card payments.

It is precisely this kind of result that drives the group. "By understanding the fullness of the social and economic structures of these attacks," explains Savage, "we're in a much better position to design and employ defenses."

Federal funding agencies seem to agree. In September 2012, the National Science Foundation announced a $10 million Frontier award to fund the UC San Diego group and their partners on a new effort to develop an empirical basis for socio-economic perspectives on Internet security.

* **Click Trajectories: End-to-End Analysis of the Spam Value Chain**, Kirill Levchenko, Andreas Pitsillidis, Neha Chachra, Brandon Enright, Márk Félegyházi, Chris Grier, Tristan Halvorson, Chris Kanich, Christian Kreibich, He Liu, Damon McCoy, Nicholas Weaver, Vern Paxson, Geoffrey M. Voelker, and Stefan Savage, *Proceedings of the IEEE Symposium and Security and Privacy*, May 2011. http://bit.ly/k4kOWO

** **Priceless: The Role of Payments in Abuse-Advertised Goods**, Damon McCoy, Hitesh Dharmdasani, Christian Kreibich, Geoffrey M. Voelker, and Stefan Savage, *Proc. ACM Conference on Computer and Communications Security*, October 2012. http://bit.ly/SdK4qg

the United States Council for Automotive Research (USCAR). The reports also caught the attention of the Department of Transportation, and played a role in the decision by automobile manufacturers to invest in new security features. Indeed, one manufacturer disclosed to *MIT Technology Review* an order-of-magnitude increase in security hiring.

Perhaps most significantly for CSE, students involved in the research have parlayed their tremendous efforts into jobs for themselves, with Stephen Checkoway joining the faculty at Johns Hopkins University, and Damon McCoy at George Mason University.

* **Comprehensive Experimental Analyses of Automotive Attack Surfaces**, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, S. Savage, K. Koscher, A. Czeskis, F. Roesner, and T. Kohno. *USENIX Security Symposium*, August 2011. http://bit.ly/V0K0lv

Experimental Security Analysis of a Modern Automobile, K. Koscher, A. Czeskis, F. Roesner, S. Patel, T. Kohno, S. Checkoway, D. McCoy, B. Kantor, D. Anderson, H. Shacham, and S. Savage. *IEEE Security and Privacy*, 2010. http://bit.ly/a3VJGl

To test the electronic control unit (ECU) in a controlled environment, the researchers immobilized the car on jack stands while mounting attacks.

CSE's 'Broad and Deep Curriculum'

The Computer Science and Engineering department prepares its undergraduate students for top placement in a competitive, high-tech job market, and for advanced studies in graduate school. From processing images and sounds in the introductory computer science course as a freshman (CSE 8A), programming robots and smartphones as a sophomore (CSE 87), rendering photorealistic images using computer graphics as a junior (CSE 168), to designing and implementing 3D, networked multiplayer games as a senior (CSE 125), our majors experience a broad and deep curriculum combining core requirements and electives in leading-edge areas as well as senior project courses. Beyond the classroom, the department has a long tradition of providing students with science, engineering and technology opportunities outside of the department. Below are just a few of the research centers and programs that are unique to UC San Diego.

DRIVING INNOVATION

The California Institute for Telecommunications and Information Technology (Calit2) is a large, multi-disciplinary research institute located next to the CSE building. Directed by CSE Prof. Larry Smarr (pictured at left), Calit2 is addressing large-scale societal issues by bringing together multi-disciplinary teams in fundamentally new ways. Wireless networks for disaster and crisis response, information technology to support digitally enabled medicine, and the creation of new media via research in computer games and novel visualization environments are just a sampling of its many research strengths. CSE undergraduates play important roles in many Calit2 research projects, either as student workers during the academic year or through the Calit2 Summer Undergraduate Research Scholars program.

www.calit2.net

SUPERCOMPUTING AT VAST SCALES

The San Diego Supercomputer Center (SDSC) is a pioneer in the cyberinfrastructure revolution. It houses a unique combination of equipment, resources and expertise for major projects in science and engineering that require complex information and coordinated technologies to make new discoveries. From the Triton Resource, a massive data-analysis and preservation system, to Dash, a unique high-performance computing system for data-intensive supercomputing, SDSC supports science at unprecedented scales. CSE undergraduates participate in SDSC in a variety of ways, including courses on parallel computing, undergraduate research projects for degree credit, and paid internships on research projects.

www.sdsc.edu

EXPLORING WITH NATIONAL GEOGRAPHIC

Co-directed by CSE Prof. Ryan Kastner, the UCSD-National Geographic Engineers for Exploration program is a unique partnership between the campus and the National Geographic Society. It provides resources, opportunity, and a platform for students to engage in real-world engineering challenges that directly impact the world of exploration. Undergraduate students work in teams together with NGS engineers and explorers to implement and deploy tools for National Geographic expeditions and conservation projects. Active projects range from developing underwater, remote-operated vehicles for exploring the Hoyo Negro cave system in Mexico, to 'camera traps' that capture images of animals in the wild, and aerial camera platforms (such as the one under construction, pictured at left) for photographing archaeological dig sites and excavations.

http://ngs.ucsd.edu

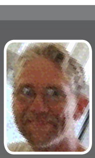
ENGINEERING IN SERVICE TO THE WORLD COMMUNITY

Founded by CSE Prof. Jeanne Ferrante, Global Teams in Engineering Service (Global TIES) is a humanitarian engineering program of the Jacobs School of Engineering. Multi-disciplinary teams of computer-science and engineering students work together with not-for-profit organizations to serve their clients in San Diego and in developing countries. CSE undergraduates have worked on dozens of projects with organizations such as the National Federation for the Blind (to develop a computer vision-based system to assist the blind with grocery shopping); Engineers Without Borders, for which a team (pictured at right) is developing solar-power devices for rural village communities in Kenya; and medical and public-health NGOs in Haiti.

http://ties.ucsd.edu

PROFESSORS (TEACHING)

Christine Alvarado's interests range from developing sketch-based user interfaces to designing a novel curriculum that makes CS more accessible and engaging. While at Harvey Mudd College she helped increase women in the CS major from 12% to 40%.



Paul Kube teaches courses in introductory object-oriented programming, object-oriented design and advanced data structures. He serves as a member of the CSE Undergraduate Committee, is Chair of UCSD's Teaching Development Advisory Committee, and is a member of the Sixth College Executive Committee.

Beth Simon works in computer science education research specializing in introductory computing experiences, multi-institutional studies, and effective pedagogies for teaching.

IN THEIR OWN WORDS

IN THEIR OWN WORDS

"The opportunities presented to me while I was a UCSD undergraduate helped me get to where I am today. The vast number of internships, research opportunities, projects and theory classes allowed for a great general education with the possibility to specialize in almost whatever you want, even video game development. I thoroughly enjoyed my time at UCSD working on my projects, and later applying and expanding on the knowledge and skills that originated in my education at UCSD."

After a summer internship at **Blizzard Entertainment** working on Diablo III, **Ryan Mourey ('12)** now works full-time at the same company as a game developer.

"I have had a great experience in the CSE department at UCSD. After becoming involved in the CSE tutoring program, I gained a new appreciation for computer science and the department as a whole. The professors are very approachable, and I intend to stay connected with the department during my graduate school career. Whether you are a tutor, research assistant, or a student in one of their classes, the faculty are readily available to help with anything."

While double-majoring in Pure Mathematics and Computer Science (with a minor in Music), **Leilani Gilpin ('11)** received an Honorable Mention for the Computing Research Association (CRA) Outstanding Undergraduate Award, and an NSF Graduate Research Fellowship. She was President of the UCSD Math Club, VP of the Triton Engineering Student Council, tutored extensively for programming courses, and competed on swimming and rowing teams. She graduated with High Honors and is now pursuing her Master's degree in Computational and Mathematical Engineering at **Stanford University**.

"The large project courses like CSE 110 and 125 gave me experience working with large groups, and provided me with the practical knowledge needed in software development. In addition, the CSE department helped me professionally with formal career fairs as well as more informal resources. Many professors were excited to point me in the right direction when it came time to interview for internships and full-time positions."

David Watson ('11) spent a summer at Apple and also multiple years as an undergraduate tutor in CSE. He is a Software Engineer at **Google**.

"The thing I liked most about CSE was the collaborative, friendly environment. Doing well in classes never felt like a competition. I really enjoyed tutoring for three years; it's a great way to help other students in the department and to work with professors. Before I came to UCSD I was very worried that I wouldn't like computer science. I had little prior experience, and I didn't really know what to expect. Looking back now, I can honestly say majoring in CSE at UCSD was one of the best decisions I've ever made."

Jennifer Chandler ('11) was a tireless CSE undergraduate tutor who, during summers, interned at Intuit and Microsoft. She received an NSF Graduate Research Fellowship and is pursuing her Ph.D. in graphics and visualization at **UC Davis**.

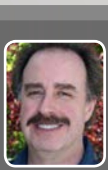
"The quality of the CSE department's tutoring program is unparalleled. The level of individual attention for students in introductory classes is phenomenal, and it breeds a culture of students giving back to their fellow students."

At UCSD, **Elliott Slaughter** ('11) received the Kenneth Bowles Scholarship for Computer Science and Engineering and the Klara D. Eckart Scholarship for computation, mathematics and physics. He was also deeply involved in CSE departmental activities, competed at two ACM International Collegiate Programming Contest World Finals. Upon graduation, Elliott received CSE's Outstanding Student Leader Award, and he is pursuing a Ph.D. in programming languages and compilers at **Stanford University**.

"The CSE department at UCSD and the Jacobs School of Engineering as a whole have given me a great college experience. The education I received was top-notch, and I was able to work on various projects and research in the department. I had many opportunities to exercise leadership and be part of something bigger, an experience that will guide me for many years."

Justin Huang ('12) was heavily involved in student organizations, eventually serving as President of the Triton Engineering Student Council and as an officer in the Computer Science and Engineering Society. He spent his summers interning at Amazon, Google, and Facebook, and is now a Ph.D. student at the **University of Washington**.

LECTURERS



Gary Gillespie emphasizes building problem-solving skills and designing software systems via insights into data structures, memory management, and object-oriented design. He brings practical experience into the classroom from years in industry as a Senior Software Engineer. He runs the undergraduate tutor training program.

Susan Marx is a full-time lecturer designing, developing and instructing introductory courses in computers and programming focused on Java, C, Information Technology, and MATLAB. Susan served as the Faculty Co-chair of the Chancellor's Advisory Committee on the Status of Women from 2010-2012.

Richard C. Ord is a full-time lecturer teaching a wide variety of CSE courses. Recent teaching awards include CSE Teacher of the Year (2003 - 2006) and Teacher of the Year at Warren College (2004) and Sixth College (2008).

UNDERGRADUATE ADMISSIONS

CSE welcomes applications to its Bachelor's degree programs in computer science, computer engineering, and bioinformatics from both incoming freshmen and transfer students. For detailed information, visit www.cse.ucsd.edu/ugrad or email questions to ugradinfo@cs.ucsd.edu.

Seeing and Seizing the Upside in Computer Vision

Boris Babenko is relatively fresh from earning his Ph.D., and he has already built and sold his first company, Anchovi Labs, to Dropbox.

Babenko earned his B.S., M.S. and Ph.D. degrees in Computer Science and Engineering at UC San Diego, the latter in the field of computer vision, under advisor Serge Belongie. "Boris has a remarkable, world-class combination of academic and entrepreneurial skill," says Belongie. "He is a source of inspiration to me and his students, and through Anchovi's success story, he sets an impressive example for a new generation of students in machine learning and computer vision who aspire to do great work with real-world impact."

Ph.D. student Boris Babenko performs at the 2010 Jacobs School of Rock, an annual event where musicians affiliated with the school – including his thesis advisor Serge Belongie – perform at Porter's Pub on the UC San Diego campus.

In his final year of graduate school, Babenko was one of 15 North American recipients of Google Ph.D. Fellowships, designed to help grad students focus on research as they complete their dissertations.

The previous summer Babenko received an internship in Google's Santa Monica research facility to work on Google Goggles, a mobile image-recognition app that allows users of Android and iOS smartphones to take a photo and search the Web for relevant information. It works on bar codes, landmark buildings, logos and other classes of images – including paintings in New York's Metropolitan Museum of Art, and even Sudoku puzzles (which it can help solve).

"Google has amazing development resources," says Babenko, who, as a Ph.D. student, focused on pattern-recognition problems that usually require a large volume of precisely labeled data. Teaching a computer system to recognize a person's face, for example, requires that the computer be exposed to many examples of that person's face. Collecting large amounts of 'training' data is time consuming, and for some applications, it can become prohibitively expensive.

Babenko's approach uses labels that are less precise and therefore less expensive to obtain. With these co-called 'weakly supervised learning' systems, training a computer to recognize a particular face might simply require photos of the person – without having to specify exactly where the person's face is in each image. This problem is more challenging due to ambiguity in location, especially when the photos contain multiple people. Yet this approach can be used to tackle a wide range of computer vision problems, and Babenko was able to apply similar ideas to recognizing landmarks, detecting pedestrians, and tracking objects in video clips in his doctoral dissertation, "Training Discriminative Computer Vision Models with Weak Supervision."

Before receiving his Ph.D. in early 2012, Babenko and advisor Serge Belongie teamed with Caltech graduate student Peter Welinder, UCSD alumnus Piotr Dollar (then a postdoc at Caltech, now at Microsoft Research) and Caltech's Pietro Perona, to co-found Anchovi Labs, a company that developed a new way to organize, browse and share personal photos. Anchovi was able to participate in Summer@HIGHLAND, a startup incubator, as well as the NSF Innovation Corps program. The company also landed a small-business NSF development grant. The company was officially incorporated in November 2011, and ten months later, the co-founders sold Anchovi Labs to Dropbox in a deal that Bloomberg Businessweek magazine reported as offering its investors "a healthy profit."

As an undergraduate, Babenko spent more than a year as a software intern at Parity Computing, and later interned at Honda Research. He also spent a summer as an undergraduate research scholar in the California Institute for Telecommunications and Information Technology (Calit2) at UCSD.

Before setting up Anchovi, Babenko consulted briefly for a company in the medical-devices industry, BioImagene, and he remains interested in microscopy, pathology and other medical imaging applications that could benefit from the new capabilities of computer vision and pattern recognition. "While my research and expertise are in the areas of machine learning and computer vision," says Babenko, "I also have an interest in scientific image analysis applications."

While he was still at UC San Diego, he worked with a team of researchers to automate the arduous process of analyzing the vast amount of data necessary for tissue engineering. Collecting information on the formation of small, terminal branches of arteries is an important step for biomaterial remodeling to help bring blood flow to a damaged area of the human body. Until now it has been done manually, a time-consuming, meticulous effort, but at the Jacobs School's 2011 ResearchExpo, Babenko and bioengineering Ph.D. student Jessica DeQuach presented a system for automating the process of counting blood vessels in medical images.

Looking back at his higher education in CSE, Babenko recalls that he started out in bioinformatics, but it didn't take long to settle on computer vision as his chosen field. "When I first encountered computer vision, I didn't realize it had such strong links to machine learning and artificial intelligence," he says. "It was a neat combination of a lot of things I found interesting." After selling his company to Dropbox, Babenko and co-founder Peter Welinder opted to stay on with Dropbox. But with so many potential applications of computer vision still on the horizon, it's unlikely that Babenko's first successful startup will be his last.

Graduate Admissions

CSE annually seeks applications to its Ph.D. and M.S. graduate programs in computer science and computer engineering. For detailed information, visit www.cs.ucsd.edu/grad or email questions to gradinfo@cs.ucsd.edu.

RESEARCH SCIENTISTS

Yuvraj Agarwal is a member of the Microelectronic Embedded Systems Laboratory and executive director of the NSF-funded Variability Expedition. His research interests include systems and networking, embedded systems and mobile computing, including building scalable and energy-efficient systems.

Ali Irturk‘s research interests include design methods, languages and tools for embedded systems, particularly those involving reconfigurable devices such as field-programmable gate arrays, in signal processing, biomedical imaging and high-performance computing applications.

Kirill Levchenko earned his Ph.D. in computer science from UC San Diego. His research interests include Internet measurement, routing, and network security. Levchenko's work in cyber-security is centered on economically-motivated criminal activity.

George Porter's research spans distributed systems, networking and data-intensive computing, with a focus on improving data-center networks to better support data-intensive applications, and enabling developers to build, deploy and operate applications that support significantly larger data sizes at low cost and with low energy requirements.

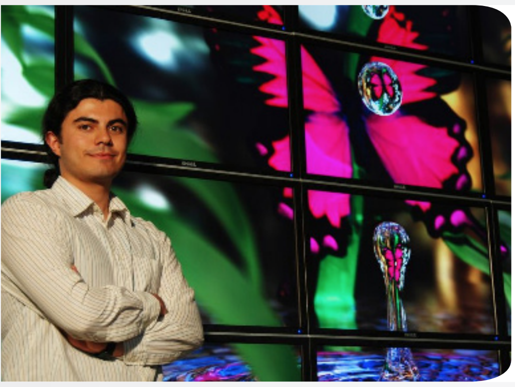
Lingjia Tang is interested in compilers, scheduling, runtime systems and novel hardware to improve performance, utilization, quality of service, predictability, reliability and energy efficiency of warehouse-scale datacenters. She also studies hardware-software co-design for better energy efficiency.

Ken Yocum does applied research at the intersection of distributed systems, networking and data processing. He is also interested in the design and construction of large, distributed systems for storing, viewing, and manipulating genomic data.

Earning a Ph.D. and Beyond: Reflections from Recent CSE Grad Students

Thomas Ristenpart (Ph.D. '10) works at the interface between systems and cryptography, two foundations for making systems and services secure in an online world where attacks are the norm. His dissertation work introduced new ways to design and analyze cryptographic hash functions, and developed format-preserving encryption techniques for legacy databases that enable services to encrypt data originally stored in plain text. Ristenpart's work also addressed security and privacy risks in emerging technologies, such as information leakage in public cloud-computing systems, and unexpected vulnerabilities in cryptographic protocols when reusing virtual machine snapshots. He is now an Assistant Professor in the Department of Computer Sciences at the **University of Wisconsin-Madison**.

"The best aspect of having trained at UC San Diego is the atmosphere of growth and improvement. The department just keeps getting better and better, which not only gives students fantastic opportunities but continuously generates an aura of excitement."



Iman Sadeghi (Ph.D. '12) developed practical approaches for the appearance modeling of specular microstructures, a key technique for realistic rendering of a variety of materials and natural phenomena such as hair, cloth and rainbows. His hair-shading model has been integrated into the production pipeline at Walt Disney Animation Studios and was used in the animated feature film *Tangled*, garnering him a film credit. And by considering the physically-based shape of water drops as well as the effect of the sun's inclination, his work was the first to present a simulation of twinned rainbows. Sadeghi is now a Software Engineer at **Google**, Los Angeles.

"I really enjoyed my time at UCSD during my graduate studies. I was able to work with many amazing professors and students during the academic year and had incredible internships in industry during summers. I was able to combine my research skills and industry experience to solve real-world problems, and that has been the most rewarding experience for me. The excellent set of faculty members, outstanding students, and the beautiful environment of San Diego make UCSD a no-brainer!"

As a basis for understanding the genome and cell processes, mass spectrometry has become a popular, efficient, high-throughput technology for studying protein expression. In her dissertation, **Natalie Castellana** (Ph.D. '12) developed techniques that significantly improve peptide identification using mass spectrometry. She developed a semi-automated pipeline that accepts mass spectra and a sequenced genome, and addresses the dual goals of annotating the genome for protein-coding genes and identifying peptide sequences in the absence of a complete, curated protein sequence database. Castellana is currently Chief Technology Officer at **Digital Proteomics**.

"It's hard to beat the CSE community at UCSD. The faculty and staff are approachable and always eager to hear new ideas about anything: research topics, computer science news, curriculum improvements, or the best restaurants in San Diego. Of course, the department wouldn't be nearly as vibrant without the world-class graduate students. In addition to being lifelong friends, my fellow graduates are now my contacts at many major universities and companies worldwide."

Small, dense, wireless sensor networks are revolutionizing our understanding of the physical world by providing fine-resolution sampling of the surrounding environment. In her dissertation, **Bridget Benson** (Ph.D. '10) developed a low-cost, underwater acoustic modem that brings this capability to underwater environments – a critical development for advancing underwater scientific and ecological analyses. Benson is now an Assistant Professor in the Electrical Engineering Department at **California Polytechnic State University San Luis Obispo**, where she continues research that spans computer engineering and aquatic sciences.

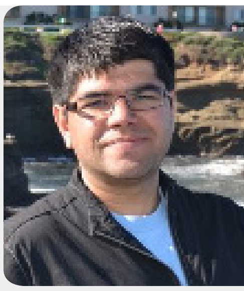
"I am very grateful for the multi-disciplinary opportunities I had as a graduate student in CSE. Not only did I work with professors and graduate students outside my concentration in CSE, but I also worked with researchers at Calit2, the Scripps Institution of Oceanography, and UCSD's Electrical and Computer Engineering department. These opportunities opened my eyes to the vast applications of computer science and engineering."

Modern software development would be impossible without compilers. Developers rely upon them for both correctness and performance. In his thesis work, **Ross Tate** (Ph.D. '12) developed rigorous techniques for advancing the state of the art in reliable compiler optimizations. Based on a new algebraic representation, his techniques can both infer and learn new opportunities for optimizations and verify the correctness of the optimizations applied by the compiler, thereby improving both program reliability and performance. Tate is now an Assistant Professor in the Computer Science Department at **Cornell University**.

"I had an excellent time at UCSD and grew a lot from the experience. Where else can I learn beach volleyball, relational algebra, and Tetris Attack all on the same day in the same department? Not to mention banter with staff, faculty, and students on the misc community mailing list. The department has a strong community culture, one I already miss, so I'm working to bring some of that to Cornell. I only hope my faculty Christmas skit can live up to the same high bar."

Meg Walraed-Sullivan (Ph.D. '12) developed techniques that improve the performance and reliability of datacenter networks – the foundations of large-scale Internet services provided by companies such as Google, Microsoft and Amazon. A key challenge is how to scale datacenter networks affordably and power-efficiently to meet the demands of a worldwide user base, without sacrificing performance. Walraed-Sullivan's dissertation developed innovative techniques for decentralized, fault-tolerant address assignment as a basis for scalable route discovery among hundreds of thousands of hosts. She is currently a postdoctoral researcher at **Microsoft Research** in Redmond, WA.

"The years I spent at UCSD were some of the best of my life. The department as a whole is incredibly supportive, with faculty who are easy to approach and always willing to engage in new projects and ideas. The sense of camaraderie among the grad students is fantastic, and inter-group collaboration is both easy and frequent. The beautiful campus, surrounding areas and of course the weather are a big bonus on top of being able to work with some of the smartest people at one of the top computer science and engineering departments."



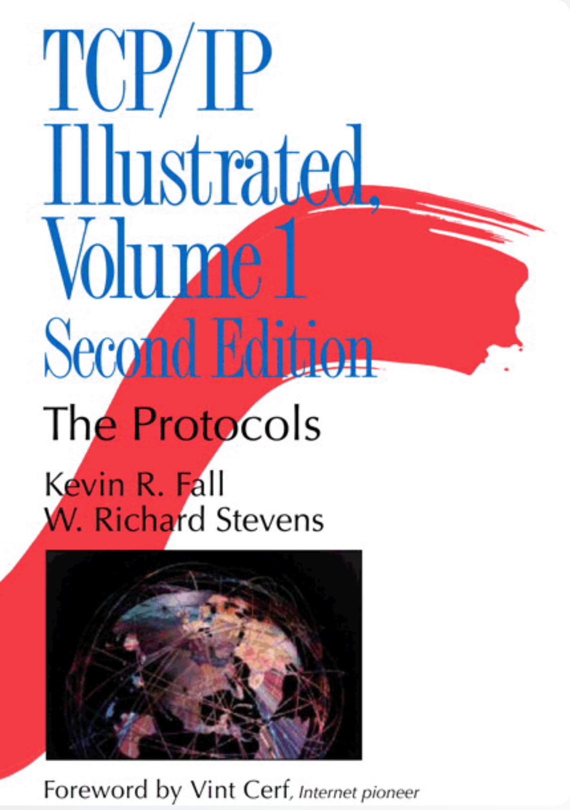
The information explosion of the last few decades has created tremendous need and opportunity to apply machine learning to large-scale data analysis. However, the sheer scale of data, whether from Internet search or human genomic databases, makes it difficult to design and analyze learning algorithms for modern, high-dimensional data. **Nakul Verma** (Ph.D. '12) developed techniques to make such machine-learning tasks tractable again. His work provided the insight and techniques for reducing the high-dimensionality of modern data sets to learn a low-dimensional 'intrinsic structure,' enabling machine-learning techniques to scale to vast amounts of data and features. Verma is now a Machine Learning Scientist at **Amazon**.

"Having done both my undergraduate and graduate studies at UCSD, I can confidently say that I could not have asked for a better place. My advisor, Sanjoy Dasgupta, encouraged me to explore and pursue my interests in machine learning. Over time, I also worked with other researchers at UCSD, gaining new insights into solving interesting and challenging problems."

Communicating in Unusual Places

There aren't very many computer scientists with a section of their CV about "Ships I have been on," but then again, CSE alumnus Kevin Fall (Ph.D., '94) has had a relatively unconventional career.

One of the ships cited on his resume, the U.S. Coast Guard Cutter Healy, is the most advanced polar icebreaker in the U.S. fleet, and Fall did two stints on board providing technical support for science cruises, including a September 2008 voyage to mark International Polar Year. Fall was also a visiting scholar at the Woods Hole Oceanographic Institution, where he investigated underwater acoustic communication and ship-to-ship *ad hoc* communication.



TCP/IP Illustrated, Volume 1 Second Edition

The Protocols

Kevin R. Fall
W. Richard Stevens

Foreword by Vint Cerf, *Internet pioneer*

"Early on I was taken with the challenges of communicating in unusual or remote places," explains Fall, who grew up in Manhattan Beach. "It began when I attended a meeting at the Jet Propulsion Laboratory on deep-space communication, and that also got me to thinking about the challenges to underwater communications and remote sites in general."

While still an undergraduate at UC Berkeley, Fall had the opportunity to work on a series of high-profile research projects. They included the DASH operating system (an early object-oriented system for symmetric multiprocessors), VorTeX (extending TeX), and the Berkeley Software Distribution of UNIX (BSD).

After Berkeley, Fall decided on UC San Diego for graduate school after working at MIT's Project Athena over the summer. He started at UCSD in 1989 working on multimedia, but quickly made the switch to operating systems under advisor and CSE Prof. Joseph Pasquale. What impressed him most at UCSD was the wide variety of academic (and not–so-academic) activities around campus. He provided tech support for the Center for Research in Language under (now Dean of Social Sciences) Jeff Elman, and later worked for the San Diego Supercomputer Center (which was still run by General Atomics) as a consultant. Fall was also involved in a five-campus University of California project called Sequoia 2000, funded by Digital Equipment Corporation, looking at how to build a better distributed computing environment for global-change researchers.

"I've always liked to dabble in multiple things," remembers Fall, who also found the time to join the dive club and qualify as a divemaster. "I ended up on the campus program review committee, which was made up of deans, the vice chancellor of academic affairs, and myself as the only grad student."

Under Pasquale, Fall did his doctoral dissertation on operating systems, specifically focusing on designing input/output (I/O) subsystems for I/O-intensive (especially multimedia) applications. "At that time, computers were so slow that the amount of data for multimedia was just a lot of overhead," he explains. "My solution was a way to avoid moving the data so much. If you wanted to move data from the network to your screen, you could connect the parts directly."

After finishing his Ph.D. in late 1994, Fall alternated as a postdoctoral researcher at UCSD and MIT, before joining Lawrence Berkeley National Laboratory to do network research. Two years later he became an adjunct professor at Berkeley, and with the dot-com boom in full force, Fall helped to get Netboost, Inc., off the ground.

"Netboost was hardware and software for something that manipulates network packets, and we had to figure out what that 'something' would be," says Fall. "In the end, we think we invented what later became known as the network processor. The major application wound up being intrusion detection."

Two years later – just before the dot-com collapse – Netboost was acquired by Intel, and the startup co-founder became an Intel employee, despite making a handsome profit in the takeover. Laughs Fall: "It was a lot of money at the time, but we were living in the Bay Area so it wasn't enough to retire on."

At Intel, Fall joined a small group of researchers who opened an Intel 'lablet' near UC Berkeley focused on 'communication in the extreme.' "I ended up working on networking in strange places," says Fall, who is credited with coining the term 'delay-tolerant networking' to describe architectures of networks that may lack continuous network connectivity, e.g., in mobile or extreme terrestrial environments.

After more than a decade, Intel decided to close its lablets, and Fall returned to his roots and interviewed with Qualcomm, Inc. He got the job, but not in their San Diego headquarters. Instead, he became principal engineer in the company's research unit in Berkeley, Calif. (where he had previously worked for Intel).

Before assuming his post at Qualcomm, Fall finished writing a book that took him more than five years to complete. He updated the bible of TCP/IP for networking professionals. "TCP/IP Illustrated, Volume 1: The Protocols, 2nd Edition" (Addison-Wesley, November 2011) was co-authored with W. Richard Stevens (but Stevens passed away before Fall began work on the new edition).

Fall was elected a Fellow of the IEEE in 2009, and he has been awarded nine U.S. patents on technologies ranging from a programmable system for processing a partitioned network infrastructure, to a compiler for a computer-programming language, including an instruction statement for handling network packets. He also just finished up four years with the U.S. Air Force's Scientific Advisory Board.

"Kevin Fall has had a really interesting career – doing just about every job that someone with a Ph.D. might do," observes CSE Prof. Stefan Savage, director of UC San Diego's Center for Networked Systems.

"The quality of instruction in CSE was great and it equipped me well for what I needed to do," observes Fall. "I also like the exposure to research and activities across the campus and beyond. The Sequoia 2000 project gave me exposure to Berkeley and Santa Barbara, and the work I did at SDSC to Los Alamos, JPL and Caltech. I know that Joe Pasquale always encouraged getting through the graduate student program basics in order to focus on my dissertation as quickly as possible, yet my various side activities wound up being nearly as important to me."

Since he now works for Qualcomm, Fall has the opportunity to be in San Diego at least once a month, and occasionally drops by the campus to meet with Joe Pasquale, lecturer Rick Ord and other friends and colleagues from grad school in CSE.

But if he ever needs a reminder of his years at UCSD, he doesn't have to look very far: to his wife Vicki. Vicki Fall (née Loeffler) was an administrative assistant in the CSE department, and they married just after he graduated.

One Alumnus' Journey into Rendering the Entire World in 3D

When Steve Rotenberg began studying Computer Science at UCSD in 1989, the department was celebrating just its second birthday. Since then he has pursued an exciting career in the computer game and 3D modeling industries, while maintaining close ties with the department.

As a senior in 1992, he started an internship at Angel Studios (later Rockstar San Diego) where, over the course of a decade, he worked as Director of Software. Among other animation and rendering projects, he developed game titles such as *Midtown Madness*, one of the first open-world racing games, and *Savage Quest*, where players hunt as a Tyrannosaurus Rex. "In the games we developed," Rotenberg recalls, "we placed a strong emphasis on pioneering simulation and rendering techniques, such as driving and collision physics for cars, and ragdoll animation for characters, to increase realism and improve the gaming experience. The technical challenge, of course, was achieving these goals on woefully underpowered PCs at the time."

After Rockstar acquired Angel Studios, Rotenberg embarked on his own entrepreneurial quest. After privately developing the core technology, he founded PixelActive as CEO and Chief Scientist in 2006 to commercialize what became known as CityScape. As a tool for the rapid prototyping, rendering and simulation of urban landscapes, CityScape quickly attracted customers ranging from Boeing, GM, and Lockheed Martin to the U.S. Army and Marine Corps. Eventually PixelActive found itself working most closely with NAVTEQ, a major provider of electronic navigable maps for GPS devices and

How CSE Helped Propel an Alumnus to eBay, Facebook and Beyond

It may not seem so unusual in a company created by twentysomethings to change the 21st century world, but Taner Halicioglu was already a veteran of several high-profile startups by the time he joined Facebook in 2004.

The company was so new that he was, he recalls, Facebook’s first ‘real’ employee (after the company’s founders).

Halicioglu, who moved back to the Bay Area after graduating from UC San Diego with a B.S. in Computer Science (Revelle College, Class of ’96), helped transform Facebook into the ubiquitous social network it quickly became. As the company’s first operations honcho on the production side, Halicioglu was responsible for building out the initial hardware infrastructure, and he helped scale Facebook over 2,000-fold in the ensuing five years.

Facebook, however, isn’t the only mega-hit on Halicioglu’s resume. Prior to joining Facebook, he worked for nearly three years at the similarly iconic eBay (2002-04). There he built monitoring tools for eBay’s Network Operations Center. The tools, many of which are still in use today, allowed eBay to keep close tabs on the performance of its databases, Web servers and applications servers, all of which were at the heart of the company’s success.

After graduating in 1996, Halicioglu spent three years at a company that, through a series of buyouts, became another hot property of the 1990s technology boom: Global Crossing. He then worked for a startup called Loudcloud (now part of HP), where he built automated tools for DNS control by systems administrators, until he joined eBay.

CSE alumnus Taner Halicioglu, Class of ’96, with girlfriend Victoria Brown at a UCSD campus event honoring alumni donors.

After his five-year run at Facebook, Halicioglu left the company to work in the Irvine headquarters of Blizzard Entertainment. He was the lead reliability engineer for its Battle.net online gaming platform, which supports millions of members worldwide who play epic, multiplayer fantasy games, including *World of WarCraft*, *StarCraft II*, and *Diablo III*.

Since leaving Blizzard in late 2011, Halicioglu has focused on investing. On his own Facebook page, the UCSD alum briefly recorded a ‘life event’ in his profile on May 18, 2012 – the day Facebook went public on the Nasdaq Stock Exchange.

Halicioglu considers himself “a computer geek, car geek, music geek, and geek geek.” He has composed and played music using MIDI synthesizers since high school in Palo Alto, Calif., not far from where he worked at Facebook. Music was also important to Halicioglu at UCSD, where his area of focus was music technology.

In his final quarter at UCSD, Halicioglu was a teaching assistant for the UNIX lab course (CSE 80), helping students learn about shells and shell scripting. Graduation in December 1996 also ended a two-year, part-time stint at the San Diego Supercomputer Center as a systems administrator for a deployment of nearly 100 workstations.

So did CSE and UCSD help shape Halicioglu’s success in California technology companies?

“UCSD helped nurture and hone some of my abilities,” says Halicioglu. “I had access to some amazing resources in both the department and at SDSC, and that’s why now I want to teach. On my own, whatever knowledge and skills I have can do only so much good. It would be much better if I could pay it forward by teaching others what I know, so they can help all sorts of people and build neat, cool things -- or even invent the ‘next big thing.’”

Halicioglu is planning to begin with a CSE 190 lecture series on operations in 2013, and he is already giving back to his alma mater in other ways. He is a member of Chancellor’s Associates and gives to the Jacobs School of Engineering’s annual fund. Halicioglu was also an inaugural member of the Jacobs School Alumni Council and the CSE Alumni Board, where he remains an active member.

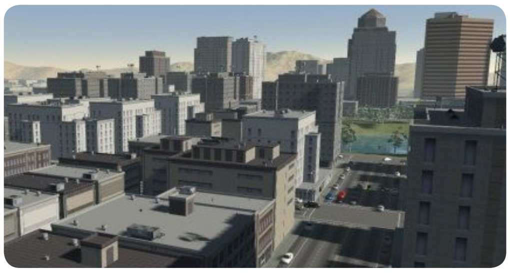
In 2009 Halicioglu picked up the tab to send a team of CSE students to the IBM-sponsored world programming competition finals in Sweden. He has also supported students in the UCSD-National Geographic Engineers for Exploration program and, most recently, he helped fund renovations in the CSE Building. The remodeling create lab spaces for the new Moxie Center, where student engineering groups can undertake hands-on projects and competitions beyond the regular CSE curriculum.

“I think it’s very important for all of our futures that the people entering the workforce today are as well, if not better trained and educated than the previous generation,” says Halicioglu, who is now back living in San Diego and began teaching CSE 191 on Computer Operations and Production Engineering during winter 2013.

“Technology is ever-changing and ever-evolving, and being able to help students explore it and get excited about it just makes me feel good!”

Web services. In 2010 NAVTEQ acquired PixelActive and both were fully merged into wireless telecom giant Nokia in 2011.

Now the Director of 3D Technology at NAVTEQ/Nokia, Rotenberg and his team are building the platform and tools for accurately modeling and beautifully rendering the entire world in 3D. “Based on this technology, people around the world will soon be able to interact with this revolutionary ‘WorldScape’ in next-generation GPS devices and smartphones,” he says.



Throughout his career, Rotenberg has remained closely involved with the department, teaching numerous courses on Computer Graphics, Computer Animation, and Video Game Programming. “Teaching at UCSD has been even more fun and rewarding than I expected,” he muses. “The students are fantastic, and being involved in the department, with its world-class graphics and computer vision groups, provides an ongoing learning experience for me as well.” Students have also benefited, and not just from the popular courses he teaches. Nearly his entire engineering team at PixelActive were fellow alumni of the Computer Science and Engineering department at UCSD.

UCSD COMPUTER SCIENCE & ENGINEERING

Jacobs School of Engineering
University of California, San Diego
9500 Gilman Drive, Dept 0404
La Jolla, CA 92093-0404
858.822.5198

www.cs.ucsd.edu