the newsletter of the International Computer Science Institute

The ICSI GAZETTE

volume nine | issue two | march 2011

featured icsi research: tailoring internet security

With relentlessly growing Internet traffic, it has never been more difficult to detect and stop malicious network incursions. Networking Group researcher Robin Sommer works to secure networks from attacks – a concern that President Obama recently described as among the "the most serious economic and national security challenges we face as a nation."

Sommer's work spans the gap between academic research and the protection of real computer networks through programs like Bro, an open-source network monitoring framework that helps defend networks against attacks.

BRO

Bro was originally developed by Networking Group researcher Vern Paxson, and the project is now led jointly by Paxson and Sommer. Bro provides users with a custom domain-specific scripting language to express their local monitoring policy. It differs from other systems in that it performs deep, semantic analysis and examines the development of a network over time – for example, tracking the Web sites a particular host has contacted. Bro is not restricted to any particular analysis approach, as most other systems are. Traditionally, signature-based systems compare observed activity against a set of low-level patterns known to indicate malicious activity; and anomaly detection systems compare new activity against an automatically learned profile of benign traffic, flagging what does not match as potentially malicious. Such systems have difficulty protecting large networks, in which traffic is diverse and the characteristics of both attacks and normal activity are constantly changing. Bro addresses this challenge by enabling users to tailor its analysis to the specifics of the local environment.

Bro has been in use at the Lawrence Berkeley National Laboratory since the late 1990s, and is now used in a growing number of networks, particularly in scientific environments. Last year, it was downloaded by about 5,000 unique IP addresses, and it now monitors networks at major universities, large research labs, supercomputing centers, and open-science communities. Many of these networks have tens of thousands of systems each – and some have as many as 100,000.

PUBLICATION AND PRACTICE

But, Sommer points out, often there is a gap between the systems that are effective in large, practical settings and the findings that are published in academic papers. Sommer believes this may be because many research groups lack ties to practical operations – at universities, network administration "is usually a totally different part of the university, and it can be challenging to build a fruitful relationship." Collaborations between researchers and operations take time to develop, and researchers sometimes underestimate the impact of operational reality on their work.

"Generally, I'm always trying to bridge that gap between laboratory research and real-world operations," he says.

As an undergraduate at the University of Paderborn in Germany, Sommer worked as a systems administrator in the computer science department. He then worked in Professor Anja Feldmann's networking research group, initially at the University of Saarbrücken and later at the Technical University of Munich, where Sommer and his colleagues collaborated closely with network administrators, who provided context about the network's traffic and crucial feedback on their research. In exchange, the group helped improve security.

"There is a give and take that is important," Sommer said. "That usually works well for both sides."

Sommer has continued to work with operations departments while at ICSI, where Networking Group researchers led by Paxson work closely with systems administrators to monitor the Institute's network. In addition, Paxon's long-standing ties to operations at the

continued on page 4

in this issue:

as i see it by Nelson Morgan, Director

"Well, in our country," said Alice, still panting a little, "you'd generally get to somewhere else – if you ran very fast for a long time, as we've been doing."

"A slow sort of country!" said the Queen. "Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast."

- *C. Dodgson (a.k.a. Lewis Carroll)*

Sometimes it seems as if the Red Queen had it right in "Through the Looking Glass"; for instance, technology moves so quickly that just keeping up takes a great effort (unless you're 15, in which case it's effortless). And yet my experience has always been that it feels like progress is extremely slow when you are working within any particular technology. You may see what needs to be done very quickly (and perhaps code up the key elements in a busy weekend), but then come all the necessary details....

The impact of technology on society seems to operate in still a different time scale – slow as molasses and then suddenly, dramatically fast. Smartphones have been around since the 1990s, but then with the introduction of the iPhone, something reached a "tipping point," and sales and impact took off. I'm not qualified to say whether the recent upheavals in the Middle East, particularly in Egypt, owed much to social networking tools, but clearly they were extensively used. There was even apparently an intervention enabled by a technology I've been interested in for a few decades, automatic speech recognition. When the Egyptian government brought down the Internet, protesters were able to call a phone number that used Google tools (including speech recognition) to generate Twitter tweets, keeping the movement connected. Maybe I'm oblivious, but Twitter just seemed like entertainment to me before this event – now it seems like it can be an enabler of social change.

This "tipping point" property, popularized by Malcolm Gladwell, seems to apply to much of the progress that we see in computer science and related disciplines. Sometimes this is due to sociological phenomena, which can suddenly make particular technologies much more important; but sometimes it is really due to discoveries that drastically shift the research landscape. The latter may not be noticeable in the public sphere for some time, but in retrospect we understand how each basic advance has changed the nature of our inquiry. For instance, hidden Markov models (HMMs) and the fundamental associated mathematics were developed in the late 1960s at IDA in Princeton, and in the early to mid-1970s these methods were applied to speech recognition at CMU and IBM. It was not until the mid- to late 1980s that HMMs were almost universally adopted for speech recognition, but in retrospect the earlier work fundamentally changed research and development for this topic. And of course, here in the Bay Area we're aware of a more fundamental part of nature that has this property – tectonic plates pushing up against one another for months, years, or even decades, until finally there is a significant earthquake.

Once a game-changing event happens in research, how do we make progress? Much of what follows is necessary but (some would say) uninspiring incremental progress. As researchers gain understanding about the basic ideas, they begin to extend it and apply it. As computational resources expand, the opportunities for exploitation of the grand ideas expand as well. Eventually the big new idea becomes the standard approach, and as such needs to be questioned and might be ultimately supplanted by the next big new idea. It's particularly the role of the young to question any standard wisdom, and it's equally the responsibility of senior researchers to push back, since most candidates for the next big idea will fail; but some of them won't, and that's the great thing to look for.

How do we get to the next set of big ideas? Moving quickly is necessary but not sufficient; like Alice, you could end up in a place that looks an awful lot like where you started. I know of no foolproof recipe for research breakthroughs, but certainly success is often associated with an interest in really understanding what is going on and a willingness to forego short-term gains. Serendipity also plays an important role, to be sure, but investigation driven by curiosity is still one of the best ways to be prepared for unexpected discoveries. Although this concept tends to be unpopular with sponsors, there is plenty of evidence that investigator-driven and curiosity-driven research is where most "next big things" come from. Since we also have to satisfy the shorter-term goals of sponsors in order to keep our work going, finding common ground between the aims of sponsor and researcher is a critical challenge. This issue of the Gazette will provide a few examples of how ICSI's researchers, past and present, have met this challenge.

news briefs

ICSI and Japan's **National Institute of Informatics** will collaborate on work in networking, computer vision, and algorithms. In a memorandum of understanding signed in January, ICSI and NII outlined their commitment to establishing a joint research program.

ICSI Director Nelson Morgan and Staff Scientist Jim Hieronymus at ICSI's BEARS Open House 2011

For the second time in the two years the contest has been held, Speech Group researchers were finalists at the **ACM Multimedia Grand Challenge**, sponsored by companies like Google and Yahoo!. Speech researchers **Gerald Friedland** and **Oriol Vinyals**, and Professor Ruzena Bajcsy and Eladio Martin of UC Berkeley were recognized for developing the first application for Android smart phones that uses WiFi and audio to accurately estimate where the phone user is inside a building. Their project, Using Android and Indoor Localization for Diaries, was one of two finalists in the Google Diaries competition. Friedland, with other ICSI researchers, won Yahoo!'s 2009 challenge with Joke-O-Mat. He will also co-chair the Grand Challenge in 2011.

Professor **Susanne Hambrusch**, a former ICSI visiting scholar, has been appointed the director of NSF's Computing and Communication Foundation (CCF), a division of Computer and Information Science and Engineering. Hambrusch, who has been on the faculty of Purdue University since 1982, was an early visitor to the Algorithms Group under Professor Richard Karp.

ICSI has had the privilege of overseeing the doctoral thesis of **Howard Lei** of the Speech Group, who graduated at the end of December.

ICSI held its annual **Open House** in conjunction with UC Berkeley's EECS Annual Research Symposium on February 17. Professor **Vern Paxson**, a senior Networking Group researcher, gave a talk about recent efforts to measure the economics of Internet attacks; **Collin Baker** led a demonstration of the latest improvements to the FrameNet Project; Networking researchers presented **Netalyzr**, a system that analyzes the extent to which Internet service providers interfere with their customers' traffic; and Speech researchers demonstrated **Joke-O-Mat**, which automatically parses television sitcom episodes, and the **Meeting Diarist**, a speaker diarization tool.

ICSI recently welcomed three new arrivals: Mona Sophia, daughter of **Gerald Friedland** and his wife Yvonne, born November 26; Linnea Viktoria, daughter of Professor **Trevor Darrell** and his wife Lisa Hagstrom, born February 3; and Asher Zev, son of **Luke Gottlieb** and his wife Emily-Rose, born March 4.

Mona Sophia Friedland

Speech Group researcher **Andreas Stolcke** was named a 2011 IEEE Fellow. Stolcke was honored for his contributions to statistical language modeling, automatic speech recognition and understanding, and automatic speaker recognition. The grade of Fellow is IEEE's highest membership status, conferred on only one-tenth of one percent of members by the IEEE Board of Directors. Fellows are recognized for their outstanding records of accomplishments.

Professor **Richard M. Karp**, leader of the Algorithms Group, was invited to give talks at three prestigious research institutes in India in January. He was a keynote speaker at the 2011 Microsoft Research School on Approximability, hosted at the Indian Institute of Science in Bangalore. He was also selected to be this year's speaker at an annual lecture in honor of Hari Sahasrabuddhe, a founding faculty member of the computer science and engineering department of the Indian Institute of Technology in Kanpur. In addition, he was invited to speak at the InfoSys Research Lab, also in Bangalore.

Oriol Vinyals of the Speech Group has received a Microsoft Research Fellowship to continue work in tele-immersion using analysis of both human speech and vision. Vinyals, one of twelve PhD students in the United States to receive the fellowship in 2011 and the only one from UC Berkeley, will help develop systems that are able to interact with humans by analyzing human gestures, gazes, and spoken questions. The work has the potential to help doctors interact with patients who are hundreds of miles away. The two-year fellowship begins this fall.

Speech Group alum **Eric Fosler-Lussier**, along with Jeremy Morris of Ohio State University, has won a 2010 best paper award from the IEEE Signal Processing Society. They were one of six teams recognized for the exceptional merit of their papers. Fosler-Lussier was a graduate student under Speech Group leader Nelson Morgan and later a postdoctoral fellow at ICSI.

continued on page 8

featured research: continued

continued from page 1

Lawrence Berkeley National Laboratory have been an asset for the group – Sommer works with the Lab's cyber-security group on a daily basis. Bro monitors the Lab's network, allowing the group to use the data gathered from monitoring for research and to try out improvements to the system.

The three-year, $3 million NSF project will work to provide Bro users with easier access to the system's capabilities and a detailed documentation.

IMPROVING PERFORMANCE

For example, the lab was the first to use Bro on a cluster of computers to monitor its network traffic. A custom high-performance frontend system looks at all network traffic and then divides it across several standard PCs so that a large volume of traffic can be analyzed simultaneously. The process allows networks with tens of thousands of systems to be monitored by inexpensive machines.

In order to divide traffic without sacrificing performance, the researchers had to find a way to flag attacks that span more than one machine. In other words, each machine in the cluster has to share not just alerts of attacks, but also the underlying analysis leading to the alert.

The cluster approach was first proposed in 2007 by Sommer and several colleagues from ICSI, LBNL, and the Technical University of Munich. It has since become a standard method for network intrusion detection systems, like Bro, in research environments.

The next step for Sommer is to make Bro more user-friendly. "[Bro] is turning into a product in that it's becoming much more widely used," he said.

But Bro is decidedly not a commodity product; it was designed as a research platform. Users need to have a high level of technical expertise in order to define their policies in Bro's scripting language – and widespread use can become problematic to support for a small group from a non-profit research institute. "We had trouble keeping up with users' demands," said Sommer.

He recently began working with the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign to increase Bro's capacity and make it more user-friendly. NCSA plans to use Bro to monitor its Blue Waters supercomputer, a machine that will be able to perform 10 quadrillion calculations every second and that will be housed in an 80,000 square foot building in Champaign, Illinois.

The three-year, $3 million NSF project will work to provide Bro users with easier access to the system's capabilities and a detailed documentation. The work with NCSA is a major step forward, Sommer says: "For the first time, we have resources for engineering that go beyond our core research projects."

HILTI

Sommer is also leading an effort to simplify the implementation of network security programs, such as firewalls or intrusion detection systems. To build such a program, one often ends up writing a lot of low-level code that is difficult to get right – code that has already been written many times for other applications. "We keep reinventing the wheel," Sommer says. His new work is "an attempt at bundling low-level functionality that is needed over and over again into a high-level platform others can build on."

HILTI – a "high-level intermediary language for traffic analysis" – and the accompanying process provide high-level abstractions specific to the network monitoring field, such as tailored support for managing the memory that a program uses to remember observed activity. It also aims to provide a suitable concurrency model for running an analysis simultaneously on many processors. For an application like Bro, this has the potential to replace today's cluster installations with a single-machine setup. Multicore processing will allow the different security analyses to "communicate faster and more directly" than they do in the distributed system.

GEO-TAGS AND GLOBAL INFERENCE

For the past year, Sommer has worked with Speech Group researcher Gerald Friedland to examine how much people unintentionally reveal about themselves through their online activity. In a technical report published last May, the two explained how a criminal might use geo-tags – geographical information embedded in photos by certain cameras and smartphones – to easily locate where the photos were taken

continued on page 8

featured researcher:
bonnie kirkpatrick

Research shows that by middle school, many girls have already decided against going into computer science careers. By that time, Algorithms Group student Bonnie Kirkpatrick had learned how to program calculators. She took the one programming class offered at her high school in rural Montana three times – and had taught herself most of the basics in her off-time.

This May, she will receive her PhD from UC Berkeley's computer science department, with a focus on algorithms for genetics aimed at deciphering the genetic indicators of both phylogenetic and family relationships. She will go on to a postdoctoral position at the University of British Columbia.

"I'm the wrong person to ask about why there are no women in computer science because I am .... I was always happy with the nerdy environment," she says. But, she adds, the increasingly male cultural associations with computer science may be the reason that fewer women are going into computer science now than 20 years ago, with the majority of college-educated women choosing social sciences, education, and biology instead.

According to the National Center for Women and Information Technology, only a quarter of jobs in technological fields today are held by women, down from 36 percent in 1991 – and those trends start in school. Only 13 percent of high school students who took the AP computer science test in 2009 were female, and the number of incoming university undergraduate women interested in a computer science degree has declined 79 percent over the past decade.

The gender gap has led to efforts designed to bring more women into the field. The Computer Research Association sponsors several programs for women, including Distributed Research Experiences for Undergraduates, formerly the Distributed Mentor Project (DMP). Through DMP, Kirkpatrick worked at Texas A&M for two summers for Professor Nancy Amato, who was the first to suggest that Kirkpatrick go to graduate school.

Last year, Kirkpatrick was one of 32 students in the U.S. awarded a scholarship from Google named in honor of Anita Borg, a pioneer in computer science who encouraged women to pursue careers in technology. In her application essay, Kirkpatrick pointed out computer science's need for "a diverse and cosmopolitan image."

"All the nerdy paraphernalia of computer science is male-gendered," she said – from Mountain Dew to computer games. To increase the number of women in computer science, those cultural associations need to change. She points to a 2009 study from the University of Washington that found that women were less interested in taking computer science courses when there were Star Trek posters on the wall, and more interested when there were nature posters up. Kirkpatrick said, "It's all about women being comfortable in that space" – not necessarily about overt sexism.

Kirkpatrick believes universities also need to make a greater effort to include women in marketing materials, and that popular culture needs to follow suit. Kirkpatrick says that young women are instilled early on with the attitude that "it's just not cool to be a woman computer scientist." Consider the roles women are assigned in television shows like the British *IT Crowd*, in which the female head of the IT department refuses to learn basic technological principles (such as the use of a Web browser) on the theory that doing so will make her a geek.

Despite universities' efforts to be more inclusive, women still comprise less than a quarter of full-time faculty in computer science in the U.S. – and less than 10 percent of the UC Berkeley EECS faculty. While an undergraduate at Montana State University, Kirkpatrick had no female computer science professors and just one female instructor.

continued on page 8

featured alum: ben gomes

Since he helped launch Google's Instant search feature last fall, ICSI alum Ben Gomes has been dubbed Google's "diplomat of search." Gomes, one of the few people to carry the title Google Fellow, worked at ICSI under AI group leader and UC Berkeley Professor Jerry Feldman during the early 1990s.

Born in Africa and raised in Bangalore, India, Gomes and a school friend, Krishna Bharat, learned about computers by playing around with a ZX Spectrum, an eight-bit home computer released in 1982. "We were both chemistry geeks, and we used to have contests to push the limits of this tiny computer" with very little information on how it worked, he said. He received his bachelor's degree from Case Western University in Ohio, and began graduate studies at UC Berkeley in 1990, joining ICSI in 1991 and earning his PhD in 1997.

While at ICSI, Gomes worked on the development of pSather, a programming language that enabled easy parallelization. Several students who worked on pSather and related projects have gone on to work at Google, including David Stoutamire and David Bailey. "There's a lot of stuff at Google that mirrors the range of interests in Jerry's group," Gomes said – from low-level programming to higher-level questions about human-computer interactions.

His interests spanned that range as well. Feldman said, "Ben was the only person who was active in all aspects of parallelization that were going on at the time, from neural modeling and hardware architecture to high level parallel programming and optimization. It was the best possible preparation for what he ended up doing."

Professor Srini Narayanan, who joined ICSI as a graduate student soon after Gomes and has since gone on to lead the AI Group, said Gomes helped create the collegial spirit in Feldman's group: "Everybody liked everybody, and Ben was largely responsible for that."

After leaving ICSI, Gomes was working at Sun Microsystems on the Java programming language when a classmate called him to tell him he had joined a small start-up in Mountain View. "He told me it seemed like a really good place with really good people," Gomes said.

Gomes and Bharat, his childhood friend, now work together at Google. Bharat invented Google News, and Gomes has been responsible for a range of changes to the search engine, from adding interface features such as spelling corrections in the world's major languages to improving how search results are ranked – what he calls "the bread and butter of what we do." Last year alone, Google introduced around 500 improvements and features to its search engine.

Google's newest feature, released last fall, shows users an image of pages returned by a search term. In order to show a relevant preview, Google has to break down the Web page and reassemble it to provide a snapshot of the most relevant parts – the heading and the portion of the page that contains the search terms. Instant Previews, developed by Gomes's team, are the next generation of "snippets" – Google was the first search engine to return results with search terms in context. Also last year, Gomes helped roll out instant search, a feature that updates search results as a user types.

Gomes said he and his team look at 200 different factors in figuring out how to get a user of Google to the best result as quickly as possible. "This company really has a strong sense of mission," he said.

visiting scholars

Since its inception, ICSI has had a strong international program consisting primarily of ties with specific countries. Current formal agreements exist with Brazil, Finland, Germany, and Switzerland. In addition, we often have visitors associated with specific research and projects.

AI
Sergio Guadarrama
Frank Hopfgartner (Germany)
Emanuel Kitzelmann (Germany)
Malte Schilling (Germany)
Carlos Subirats

Algorithms
Fabian Gieseke (Germany)
Oliver Kramer (Germany)
Shuai Cheng Li
Matthias Mnich (Germany)
Benjamin Satzger (Germany)
Ron Shamir

Speech
Arlo Faria (Brazil)
Joaquin Gonzalez
Bernd Meyer (Germany)
Stefan Steidl (Germany)

Architecture
Paula Herber (Germany)

Networking
Somaya Arianfar (Finland)
Joos-Hendrik Böse (Germany)
Gregor Maier (Germany)
Dmitriy Kuptsov (Finland)
Matti Mantere (Finland)
Boris Nechaev (Finland)
Petri Savolainen (Finland)
Baohua Yang

Vision
Peer Stelldinger (Germany)

Campus Affiliation / Other
Christoph Goebel (Germany)
Kai Huotari (Finland)
Tommi Lampikoski (Finland)
Nils Peters (Germany)


Arlo Faria


Matthias Mnich


Paula Herber


Matti Mantere

FEATURE

continued from page 4

with startling accuracy, as close as a few feet. Sommer and Friedland looked at how a criminal might use information from multiple online sources – say, a photo of someone's bike posted to a classified ads site like Craigslist, along with the information that the seller is home only after 6 p.m. – to mount an attack in real life.

Sommer will continue to work with Friedland on understanding the risk of potential inference chains that correlate personal information across independent Web sites. Work is planned to incorporate speech and image recognition to link, for example, YouTube videos and online photos.

Sommer is also involved in monitoring the traffic of residential networks in Europe and rural India. He looks at broad trends like how many homes are infected and how many lines are browsing at any one time.

Monitoring networks in India has been particularly demanding because not only is the infrastructure so different from that in the U.S., but other unforeseen challenges have arisen. Internet service providers supply some customers with Internet service; these customers, in turn, supply other customers. In one area examined by Sommer and his colleagues, Internet service is occasionally interrupted by a group of monkeys that take down the line.

KIRKPATRICK

continued from page 5

Kirkpatrick's pet peeve is the use of "he" in sentences like "A programmer needs to use all the tools he has." Even at Berkeley, where there's an "underlying value of diversity... professors still assume the typical student is male." This attitude perpetuates the disparities in the field by making female computer scientists more uncomfortable and more likely to leave the field.

The good news, says Kirkpatrick, is that this attitude can be changed if universities focused more on the increasingly interdisciplinary aspects of computer science. Computer science is no longer just about programming – Kirkpatrick's work, for example, has a focus on biology.

"What we need is to change the [male cultural] associations with computer science," she said.

NEWS

continued from page 3

Shuai Cheng Li, a postdoctoral fellow in the Algorithms Group, has received the Outstanding Achievement in Graduate Studies Designation from the University of Waterloo, Canada. The distinction is given to one person from each graduate department per year.

UC Berkeley students led by Speech Group affiliate **Dan Klein** have won a competition at the AIIDE Conference to create an artificial intelligence system that plays the video game StarCraft. The Berkeley Overmind Project, which has been featured in *New Scientist* and MTV's news blog, uses a variety of artificial intelligence techniques to calculate what moves it should take to win the game.

ICSI researchers have chaired a number of program committees for international conferences in the past year. Networking researchers **Mark Allman** and **Christian Kreibich** and external fellow **Sylvia Ratnasamy** chaired program committees for the Internet Measurement Conference, the Conference on Detection of Intrusions and Malware and Vulnerability Assessment, and the Symposium on Networked Systems Design respectively. Algorithms researcher **Eran Halperin** co-chaired the population genomic program committee for the International Conference on Intelligent Systems for Molecular Biology. In addition, ICSI researchers edited proceedings of the Symposium on Recent Advances in Intrusion Detection 2010 and selected papers from the International Symposium on Multimedia 2009, published in the *International Journal on Semantic Computing* in November.

AI Group members have received a grant from the John Templeton Foundation, an organization that supports research on "big questions" in science. Professor **Jerome Feldman** and group leader Professor **Srini Narayanan** received funding to broaden applied work on the Neural Theory of Language to language communities, public discourse, and more fundamental questions. Their project is titled "Embodied Cognition, Communities, and Foundational Issues in AI and Mathematics." Feldman also presented a paper at a recent Templeton Symposium on Top-Down Causation. A version of the paper, "Causality and Context in Cognitive Science," will be published this year by the Royal Society.

publications listing

M. Allman. On Building Special-Purpose Social Networks for Emergency Communication. ACM Computer Communication Review, Vol. 40, Issue 5, October 2010.

J. Baumbach. On the Power and Limits of Evolutionary Conservation – Unraveling Bacterial Gene Regulatory Networks. Nucleic Acids Research, Vol. 38, No. 22, pp. 7877-7884, December 2010.

M. Bermudez-Edo, N. Hurtado-Torres, and J. A. Aragon-Correa. The Importance of Trusting Beliefs Linked to the Corporate Website for Diffusion of Recruiting-Related Online Innovations. Information Technology and Management, Vol. 11, No. 4, pp. 177-189, December 2010.

R. Böhme. Towards Insurable Network Architectures. Journal of Information Technology, Vol. 52, No. 5, pp. 290-293, September 2010.

R. Böhme and M. Felegyhazi. Optimal Information Security Investment with Penetration Testing. Proceedings of the Conference on Decision and Game Theory for Security (GameSec 2010), Berlin, Germany, pp. 21-37, November 2010.

S. Bozonnet, N. Evans, X. Anguera, O. Vinyals, G. Friedland, and C. Fredouille. System Output Combination for Improved Speaker Diarization. Proceedings of the 11th International Conference of the International Speech Communication Association (Interspeech 2010), Makuhari, Japan, pp. 2642-2645, September 2010.

M. Casado, T. Koponen, R. Ramanathan, and S. Shenker. Virtualizing the Network Forwarding Plane. Proceedings of the Workshop on Programmable Routers for Extensible Services of Tomorrow (PRESTO 2010), Philadelphia, Pennsylvania, December 2010.

K. Chandrasekaran, R. M. Karp, E. Moreno-Centeno, and S. Vempala. Algorithms for Implicit Hitting Set Problems. Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA 2011), San Francisco, California, January 2011.

J. Choi, A. Janin, and G. Friedland. The 2010 ICSI Video Location Estimation System. Proceedings of the MediaEval 2010 Workshop, Pisa Italy, October 2010.

M. Christoudias, R. Urtasun, M. Salzmann, and T. Darrell. Learning to Recognize Objects from Unseen Modalities. Proceedings of the 11th European Conference on Computer Vision (ECCV 2010), Crete, Greece, Part I, pp. 677-691, September 2010.

B. Doerr, T. Jansen, D. Sudholt, C. Winzen, and C. Zarges. Optimizing Monotone Functions Can Be Difficult. Proceedings of the 11th International Conference on Parallel Problem Solving from Nature (PPSN 2010), Krakow, Poland, pp. 42-51, September 2010.

M. Ferber, T. Rauber, and S. Hunold. Combining Object-Oriented Design and SOA with Remote Objects over Web Services. Proceedings of the 8th IEEE European Conference on Web Services, Ayia Napa, Cyprus, pp. 83-90, December 2010.

G. Friedland. Review of C. Müller-Tomfelder, "Tabletops - Horizontal Interactive Displays." ACM Computing Reviews, CR138453, October 2010.

G. Friedland, J. Chong, and A. Janin. Parallelizing Speaker-Attributed Speech Recognition for Meeting Browsing. Proceedings of the 2010 IEEE International Symposium on Multimedia (ISM2010), Taiwan, pp. 121-128, December 2010.

G. Friedland, J. Chong, and A. Janin. A Parallel Meeting Diarist. Proceedings of the Workshop on Searching Spontaneous Conversational Speech (SSCS) at the ACM International Conference on Multimedia (ACM Multimedia 2010), Florence, Italy, 57-60, October 2010.

G. Friedland, L. Gottlieb, and A. Janin. Narrative-Theme Navigation for Sitcoms Supported by Fan-Generated Scripts. Proceedings of the Third International Workshop on Automated Information Extraction in Media Production (AIEMPro '10) at the ACM International Conference on Multimedia (ACM Multimedia 2010), Florence, Italy, pp. 3-8, October 2010.

G. Friedland and M.-L. Shyu, eds. Selected Papers from the 11th IEEE International Symposium on Multimedia (ISM2009). International Journal on Semantic Computing, Vol. 4, No. 2, November 2010.

G. Friedland, O. Vinyals, and T. Darrell. Multimodal Location Estimation. Proceedings of the ACM International Conference on Multimedia (ACM Multimedia 2010), Florence, Italy, pp. 1245-1251, October 2010.

G. Friedland, C. Yeo, and H. Hung. Dialocalizaton: Acoustic Speaker Diarization and Visual Localization as Joint Optimization Problem. ACM Transactions on Multimedia Computing, Communications, and Applications, Vol. 6, No. 4, Article 27, November 2010.

T. Friedrich, J. He, N. Hebbinghaus, F. Neumann, and C. Witt. Approximating Covering Problems by Randomized Search Heuristics using Multi-Objective Models. Evolutionary Computation, Vol. 18, No. 4, pp. 617-633, Winter 2010, December 2010.

M. Fritz, K. Saenko, and T. Darrell. Size Matters: Metric Visual Search Constraints from Monocular Metadata. Proceedings of the 24th Annual Conference on Neural Information Processing Systems (NIPS 2010), Vancouver, Canada, pp. 622-630, December 2010.

I. Ganichev, B. Dai, B. Godfrey, and S. Shenker. YAMR: Yet Another Multipath Routing Protocol. ACM SIGCOMM Computer Communication Review, Vol. 40, Issue 5, October 2010.

D. Gillick. Can Conversational Word Usage Be Used to Predict Speaker Demographics? Proceedings of the 11th International Conference of the International Speech Communication Association (Interspeech 2010), Makuhari, Japan, September 2010.

C. Grier, K. Thomas, V. Paxson, and M. Zhang. @spam: The Underground on 140 Characters or Less. Proceedings of the 17th ACM Conference on Computer and Communications Security, Chicago, Illinois, pp. 27-37, October 2010.

publications, cont.

T. Hein and O. Kramer. Recognition and Visualization of Music Sequences Using Self-Organizing Feature Maps. Proceedings of the 33rd Annual German Conference on AI: Advances in Artificial Intelligence (KI 2010), Kralsruhe, Germany, pp. 160-167, September 2010.

S. Hunold, R. Hoffmann, and F. Suter. Jedule: A Tool for Visualizing Schedules of Parallel Applications. Proceedings of the 39th International Conference on Parallel Processing Workshops (ICPPW 2010), San Diego, California, pp. 169-178, September 2010.

D. Imseng and G. Friedland. Tuning-Robust Initialization Methods for Speaker Diarization. IEEE Transactions on Audio, Speech, and Language Processing, Vol. 18, Issue 8, pp. 2028-2037, November 2010.

P. Jain, B. Kulis, and I. Dhillon. Inductive Regularized Learning of Kernel Functions. Proceedings of the 24th Annual Conference on Neural Information Processing Systems (NIPS 2010), Vancouver, Canada, pp. 946-954, December 2010.

A. Janin, L. Gottlieb, and G. Friedland. Joke-O-Mat HD: Browsing Sitcoms with Human Derived Transcripts. Proceedings of the ACM International Conference on Multimedia (ACM Multimedia 2010), Florence, Italy, pp. 1591-1594, October 2010.

S. Jha, R. Sommer, and C. Kreibich, eds. Proceedings of the 13th International Symposium on Recent Advances in Intrusion Detection (RAID 2010), Ottawa, Canada. Springer Verlag, September 2010.

Y. Jia, M. Salzmann, and T. Darrell. Factorized Latent Spaces with Structured Sparsity. Proceedings of the 24th Annual Conference on Neural Information Processing Systems (NIPS 2010), Vancouver, Canada, pp. 982-990, December 2010.

B. Kirkpatrick. Haplotypes versus Genotypes on Pedigrees. Proceedings of the 10th Workshop on Algorithms in Bioinformatics (WABI 2010), Liverpool, United Kingdom, pp. 136-147, September 2010.

B. Kirkpatrick, H. Finucane, Y. Reshef, and R. M. Karp. Algorithms for Comparing Pedigree Graphs. Presented at the 10th Workshop on Algorithms in Bioinformatics (WABI 2010), Liverpool, United Kingdom, September 2010.

M. Knox and G. Friedland. Multimodal Speaker Diarization Using Oriented Optical Flow Histograms. Proceedings of the 11th International Conference of the International Speech Communication Association (Interspeech 2010), Makuhari, Japan, pp. 290-293, September 2010.

T. Koponen, M. Casado, N. Gude, J. Stribling, L. Poutievski, M. Zhu, R. Ramanathan, Y. Iwata, H. Inoue, T. Hama, and S. Shenker. Onix: A Distributed Control Platform for Large-scale Production Networks. Proceedings of the 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI 10), Vancouver, Canada, pp. 351-364, October 2010.

T. Kötzing, F. Neumann, D. Sudholt, and M. Wagner. Simple Max-Min Ant Systems and the Optimization of Linear Pseudo-Boolean Functions. Proceedings of the 11th Foundations of Genetic Algorithms Workshop (FOGA XI), Schwarzenberg, Austria, January 2011.

O. Kramer. Computational Intelligence and Sustainable Energy: Case Studies and Applications. ICSI Technical Report TR-10-010, November 2010.

O. Kramer. Machine Symbol Grounding and Optimization. Proceedings of the Third International Conference on Agents and Artificial Intelligence (ICAART 2011), Rome, Italy, January 2011.

C. Kreibich and M. Jahnke, eds. Proceedings of the 7th International Conference on Detection of Intrusions and Malware (DIMVA 2010), Bonn, Germany, Springer, 2010.

C. Kreibich, N. Weaver, B. Nechaev, and V. Paxson. Netalyzr: Illuminating The Edge Network. Proceedings of the Internet Measurement Conference (IMC 2010), Melbourne, Australia, pp. 246-259, November 2010.

J. Lässig and D. Sudholt. Adaptive Population Models for Offspring Populations and Parallel Evolutionary Algorithms. Proceedings of the 11th Foundations of Genetic Algorithms Workshop (FOGA XI), Schwarzenberg, Austria, January 2011.

J. Lässig and D. Sudholt. Experimental Supplements to the Theoretical Analysis of Migration in the Island Model. Proceedings of the 11th International Conference on Parallel Problem Solving from Nature (PPSN 2010), Krakow, Poland, pp. 224-233, September 2010

J. Lässig and D. Sudholt. General Scheme for Analyzing Running Times of Parallel Evolutionary Algorithms. Proceedings of the 11th International Conference on Parallel Problem Solving from Nature (PPSN 2010), Krakow, Poland, pp. 234-243, September 2010.

H. Lei. Structured Approaches to Data Selection for Speaker Recognition. UC Berkeley dissertation, December 2010.

H. Lei, J. Choi, A. Janin, and G. Friedland. Persona Linking: Matching Uploaders of Videos Across Accounts. ICSI Technical Report TR-10-009, October 20, 2010.

E. Martin, O. Vinyals, G. Friedland, and R. Bajcsy. Precise Indoor Localization Using Smart Phones. Proceedings of the ACM International Conference on Multimedia (ACM Multimedia 2010), Florence, Italy, pp. 787-790, October 2010.

C. Muthukrishnan, V. Paxson, M. Allman, and A. Akella. Using Strongly Typed Networking to Architect for Tussle. Proceedings of the 9th ACM Workshop on Hot Topics in Networks (HotNets-IX), Monterey, California, Article No. 9, October 2010.

B. Pasaniuc, R. Avinery, T. Gur, C. F. Skibola, P. M. Bracci, and E. Halperin. A Generic Coalescent-Based Framework for the Selection of a Reference Panel for Imputation. Genetic Epidemiology, Vol. 34, Issue 8, pp. 773-782, December 2010.

L. Popa, N. Egi, S. Ratnasamy, and I. Stoica. Building Extensible Networks with Rule-Based Forwarding. Proceedings of the 9th USENIX Symposium on Operating Systems Design and Implementation (OSDI 10), Vancouver, Canada, pp. 379-391, October 2010.

L. Popa, A. Ghodsi, and I. Stoica. HTTP as the Narrow Waist of the Future Internet. Proceedings of the 9th ACM Workshop on Hot Topics in Networks (HotNets-IX), Monterey, California, Article No. 6, October 2010.

L. Popa, S. Ratnasamy, G. Iannaccone, A. Krishnamurthy, and I. Stoica. A Cost Comparison of Datacenter Network Architectures. Proceedings of the 6th ACM International Conference on Emerging Networking Experiments and Technologies, Philadelphia, Pennsylvania, Article No. 16, December 2010.

L. Popa, M. Yu, S. Y. Ko, S. Ratnasamy, and I. Stoica. CloudPolice: Taking Access Control out of the Network. Proceedings of the 9th ACM Workshop on Hot Topics in Networks (HotNets-IX), Monterey, California, Article No. 7, October 2010.

M. Preuss, I. R. König, J. R. Thompson, J. Erdmann, D. Absher, T. L. Assimes, S. Blankenberg, E. Boerwinkle, L. Chen, L. A. Cupples, A. S. Hall, E. Halperin, et al. Design of the Coronary Artery Disease Genome-Wide Replication and Meta-Analysis (CARDIoGRAM) Study – A Genome-Wide Association Meta-Analysis Involving More than 22,000 Cases and 60,000 Controls. Circulation: Cardiovascular Genetics, Vol. 3, pp. 475-483, October 2010.

S. Ravuri and N. Morgan. Using Spectro-Temporal Features to Improve AFE Feature Extraction for ASR. Proceedings of the 11th International Conference of the International Speech Communication Association (Interspeech 2010), Makuhari, Japan, pp. 1181-1184, September 2010.

K. Riedhammer, B. Favre, and D. Hakkani-Tür. Long Story Short - Global Unsupervised Models for Keyphrase Based Meeting Summarization. Speech Communication, Vol. 52, Issue 10, pp. 801-815. DOI:10.1016/j.specom.2010.06.002, October 2010.

R. Ronen, I. Gan, S. Modai, A. Sukacheov, G. Dror, E. Halperin, and N. Shomron. miRNAkey: A Software for microRNA Deep Sequencing Analysis. Bioinformatics, Vol. 26, Issue 20, pp. 2615-2616, October 2010.

K. Saenko, B. Kulis, M. Fritz, and T. Darrell. Adapting Visual Category Models to New Domains. Proceedings of the 11th European Conference on Computer Vision (ECCV 2010), Crete, Greece, Part IV, pp. 213-226, September 2010.

B. Satzger, A. Pietzowski, and T. Ungerer. Autonomous and Scalable Failure Detection in Distributed Systems. International Journal of Autonomous and Adaptive Communications Systems, Vol. 4, No. 1, pp. 61-77, January 2011.

D. Sudholt. General Lower Bounds for the Running Time of Evolutionary Algorithms. Proceedings of the 11th International Conference on Parallel Problem Solving from Nature (PPSN 2010), Krakow, Poland, pp. 124-133, September 2010.

D. Sudholt. Using Markov-Chain Mixing Time Estimates for the Analysis of Ant Colony Optimization. Proceedings of the 11th Foundations of Genetic Algorithms Workshop (FOGA XI), Schwarzenberg, Austria, January 2011.

D. Sudholt and C. Zarges. Analysis of an Iterated Local Search Algorithm for Vortex Coloring. Proceedings of the 21st International Symposium on Algorithms and Computation (ISAAC 2010), Jeju Island, Korea, December 2010.

A. Thom and O. Kramer. Acceleration of DBSCAN-Based Clustering with Reduced Neighborhood Evaluations. Proceedings of the 33rd Annual German Conference on AI: Advances in Artificial Intelligence (KI 2010), Kralsruhe, Germany, pp. 195-202, September 2010.

F. Valente, M. Magimai Doss, C. Plahl, S. Ravuri, and W. Wang. A Comparative Large Scale Study of MLP Features for Mandarin ASR. Proceedings of the 11th International Conference of the International Speech Communication Association (Interspeech 2010), Makuhari, Japan, pp. 2630-2363, September 2010.

C. Vaquero, O. Vinyals, and G. Friedland. A Hybrid Approach to Online Speaker Diarization. Proceedings of the 11th International Conference of the International Speech Communication Association (Interspeech 2010), Makuhari, Japan, pp. 2642-2645, September 2010.

O. Vinyals, G. Friedland, and N. Morgan. Discriminative Training for Hierarchical Clustering in Speaker Diarization. Proceedings of the 11th International Conference of the International Speech Communication Association (Interspeech 2010), Makuhari, Japan, pp. 2326-2329, September 2010.

O. Vinyals, E. Martin, and G. Friedland. Multimodal Indoor Localization: An Audio-Wireless-Based Approach. Proceedings of the Fourth IEEE International Conference on Semantic Computing (ICSC-2010), Pittsburgh, Pennsylvania, pp. 120-125, September 2010.

H. Ziekow and O. Günther. Sharing RFID and Complex Event Data Among Organizations. Information Systems Frontiers, Vol. 12, No. 5, pp. 541-549, November 2010.

INTERNATIONAL
COMPUTER SCIENCE
INSTITUTE

THE INTERNATIONAL COMPUTER SCIENCE INSTITUTE (ICSI) is the only independent, non-profit US lab solely conducting open, non-proprietary, pre-competitive research in computer science. Affiliated with the University of California campus in Berkeley, ICSI's mission is to further research in computer science through international collaboration, and further international collaboration through research in computer science. ICSI provides a haven for computer scientists to conduct long-term research without commercial limitations and with few faculty pressures. ICSI has significant efforts in these major research areas: Internet research, including Internet architecture, related theoretical questions, and network services and applications; theoretical computer science, including applications to bioinformatics; artificial intelligence, particularly for applications to natural language understanding; natural speech processing; computer architecture; and computer vision.

The ICSI Gazette is published twice a year by the International Computer Science Institute (ICSI).

INTERNATIONAL
COMPUTER
SCIENCE
INSTITUTE

1947 Center Street Suite 600
Berkeley, CA 94704 USA
v 510 666 2900
www.icsi.berkeley.edu