

featured research: defending the internet



Christian Kreibich. Visiting scientists, student research assistants, and summer interns also contribute to ICSI security work. The team's overall goal is to make the Internet safer for all users. Progress toward this goal is made through targeted research projects, each focusing on a specific security problem. Many of these projects at ICSI are interrelated, building on each others' success, while a few look at the problem from completely new angles. This research, both past and on-going, has been supported by funds and donations from the National Science Foundation (NSF), the Department of Homeland Security (DHS), the Department of Energy (DOE), Cisco,

Intel, ESnet, Microsoft Research, VMware, and HP. The following project descriptions summarize several of the current Internet security projects at ICSI.

ICSI HONEYFARM

The ICSI Honeyfarm project was designed as a way to study the spread of malware without contributing to the infection epidemic in the real world. The Honeyfarm lures Internet attackers to its "Honeypots", computers and virtual computers designed to be vulnerable to attacks. This network of Honeypots isolates the malware, allowing researchers to study how it spreads and how it goes about causing damage under laboratory conditions. The Honeyfarm prototype that was created by Weidong Cui, a researcher at Microsoft (and this issue's featured alum), now has an entirely new implementation developed by Christian Kreibich and former intern Steve Hanna that addresses key changes in the threat landscape. One such change is that spam has become a common method of spreading

Continued on Page 4

in this issue:

page 2
As I See It
by Nelson Morgan, Director

page 3
News Briefs

page 6
Featured Alumni

page 9
Visiting Scholars

page 10
Publications

Internet security is an issue of worldwide concern, as malicious hackers continue to invent new ways to steal personal information, infect computers with viruses, and otherwise use the Internet for nefarious purposes. Scientists at ICSI are fighting back, studying how malware is created and distributed, and developing methods to keep the Internet safe in a rapidly changing environment.

Security work at ICSI includes practical applications to prevent malware from infecting computers as well as the more theoretical study of malware to better understand how it works, and thus inform the development of future security software.

The ICSI security team is led by Vern Paxson, also Associate Professor in the Electrical Engineering and Computer Science (EECS) Department at UC Berkeley and Staff Scientist at Lawrence Berkeley Laboratory (LBL). Current ICSI scientists are Mark Allman, Nicholas Weaver, Robin Sommer, and



as i see it by Nelson Morgan, Director

We need to know the candidates' qualifications for understanding and judging science, and for speaking intelligently about science and technology to the leaders of other nations in planning our collective global future.

-- Donald Kennedy, editor-in-chief of Science

As I write this we are in the heat of the Primary season for the 2008 Presidential election in the U.S. I find myself to be enthusiastic for one of the candidates for the first time in many years. No, I'm not going to say who that might be; we are, after all, a public not-for-profit corporation, and I should not advocate any particular candidate in this public forum. But what's relevant for this column and this institute is one of the requirements for a new president – having enough understanding of science and technology to put forward effective policies.

To some extent, technology moves forward on its own – innovation generates financial success, at least for some, and researchers will continue to invent regardless of (or despite) government policies. But this is a limited view of what really happens. Government policies have huge effects on scientific and technological progress. Diminished resources for schools limit the “seed corn” of innovation, and shortsighted approaches to research and development limit the potential for future advances. For the most part, industry does not (and perhaps can not, at least currently) focus on long-term research, and while there are admirable efforts to support exceptional student researchers with internships and fellowships, there are relatively few such opportunities. And of course, the core capabilities of students pursuing advanced studies are limited by the problems in K-12 education in the U.S.

Increasingly, government support for technological research has had the same kind of short-term perspective that corporations have adopted to satisfy their Boards and stockholders. Agencies like the National Science Foundation and the National Institutes of Health are explicitly designed to support longer-term studies, but only a tiny fraction of proposals to these organizations are funded. And while all of us who serve as reviewers support the notion of high risk, transformative research, the moderating effect of peer review panels tends to favor what is safe, since anything controversial is likely to be killed by at least one reviewer, given the small fraction of proposals that can be funded.

Providing more funding for research and training is not sufficient to handle these problems. Government also needs to provide effective leadership, for instance by appointing agency leadership that celebrates rationality and rewards honest scholarship, rather than censoring science for political goals. Hopefully this will be a characteristic of the next Administration, whoever takes office next January 20. But sufficient funding is definitely a necessary condition, and we're not there now. Increasing the budgets for science and technology has seemingly been a consensus view in government for some time now, but the reality has not matched the rhetoric. There are many reasons for this, and this column is not the right place to enumerate them (nor am I expert in such analysis); but moving past the current blockages to properly fund research and education is critical to the U.S. remaining a great nation.

Does this rant have much to do with ICSI? As our name indicates, we are an international institute, and receive funding from many sources outside the U.S. government. Many of the Institute researchers come from elsewhere in the world, and most will ultimately return to their home countries. And yet, we are deeply affected by government research policies. U.S. government support provides most of our funding, and the increasing emphasis on short-term deliverables for many agencies makes it much more difficult to maintain a consistent research focus for the term of a student's Ph.D. Since we are committed to continuing our long-term research in areas such as Internet security, bioinformatics, and human language technology, senior researchers must spend an increasing amount of time hunting for funding that can be pieced together into a coherent whole. We don't shy away from competition, and our “hit rate” is far greater than the national average; but the hunt for support has become a significant drag on our human resources.

Despite these concerns, ICSI thrives, in no small part due to our international program. Due to our training mission, and our success in attracting foreign visitors, we are able to continue research in our core areas despite the inevitable fluctuations in U.S. government funding. Industrial support, while not reaching the proportions that we saw during the dot-com boom, has also continued to supplement both U.S. and international funding. But we will observe the U.S. Presidential campaign with great interest, as the choices of the next administration will have enormous effects on research and education institutions throughout the country, including ICSI.

In this issue of the Gazette, we will feature an overview of ICSI work on Internet Security. Our News section will also report a great honor received by our leader in this work, Prof. Vern Paxson.

news briefs



Members of Brazilian Agency for Industrial Development

On October 30th, ICSI signed a Memo of Understanding (MOU) with the **BRAZILIAN AGENCY FOR INDUSTRIAL DEVELOPMENT**. Mr. Reginaldo Arcuri, President of the Agency, visited ICSI to sign the memo along with officials from the Brazilian Development Bank. The MOU will be the framework within which future activities between ICSI and Brazil, such as a visitor program, will be organized.

Two of ICSI's German postdoctoral fellows are working with our newest industrial partners. **GERALD FRIEDLAND**, who came to ICSI for a 2007 fellowship, extended his research visit another year with funding from the Silicon Valley company **APPSCIO, INC.** **FELIX SALFNER** arrived at ICSI in January of 2008 and is working with German company **SAP's** Palo Alto office. Salfner and Friedland also receive funding for their research visits from **DAAD** in Germany, ICSI's longest-running international visitor program sponsor.

The **2007 SPANISH CALL** for proposals was issued in January. Applications were accepted through February 24, 2008, and are now being reviewed.

The Regional Council of Tuscany has awarded the Giulio Predi Prize in Science and Democracy to Professor **GEORGE LAKOFF** of the Berkeley Linguistics Department. Lakoff is a longtime collaborator with the AI Group at ICSI. This is the first time the prize has been awarded and Professor Lakoff was the unanimous choice of the scientific committee. He will be accepting the prize in Florence, Italy on November 24, 2008.

PAUL KAY of ICSI's AI group was featured in articles in Nature News by Kerri Smith entitled [Perception coloured by language](#) and in Wired by Brandon Keim entitled [Babies See Pure Color, but Adults Peer Through Prism of Language](#) on March 3, and also in a March 10 article in Popular Science by Matt Ransford entitled [Color and Language](#). The articles describe studies by Kay and collaborators which suggest that children have categorical perception for color in the right (non-language) side of the brain; however, in adults with intact language, color categorization is found in the left side of the brain. This particularly affects colors that are easiest to name - e.g., red and blue, in comparison to pinkish-purple.

UMIT GUZ, Visiting Postdoctoral Researcher in 2007, received The Scientific and Technological Research Council of Turkey (TUBITAK) CAREER



Umit Guz

Award for his project, "Extracting and Using Prosodic Information for Turkish Spoken Language Processing". Guz has returned to Turkey where he will conduct this research over the next two years, advised by ICSI's Dilek Hakkani-Tur and SRI's Gokhan Tur and Murat Akbacak.

VERN PAXSON of the Networking Group collaborated with Carnegie Mellon and UC San Diego to design tools to fight the growth of Internet black markets. Jason Franklin, a PhD student at CMU and former ICSI visitor, said "Our research monitoring found that more than 80,000 potential credit card numbers were available through these illicit underground web economies." To read more about the Networking Group's efforts, see this issue's featured research article.

NELSON MORGAN, Director of ICSI, was quoted on [globeandmail.com](#), a national newspaper in Canada, in an article called [Reality TV: When the tube talks back](#). The article discusses the evolving ways in which people interact with technologies such as speech recognition.

Congratulations to **JISUP HONG** and his wife Sara. Their son, Jonathan Jingul Hong, was born on September 7th at 8:49 a.m. Congratulations to **BIRTE LOENNEKER RODMAN**,

Visiting Scientist with the AI Group, and her husband Matjaz Rodman on the birth of their daughter Amaia Rodman on December 2nd



Amaia Rodman

at 4:00 a.m. Also, congratulations to **LEAH HITCHCOCK-YBARRA** and her husband Chris on the birth of their son Andrew Ybarra on March 7th at 5:02 a.m. These are the first babies for each couple.

GERALD FRIEDLAND, a German postdoc working with the Speech Group, received the IEEE Computer Society Distinguished Service Award



Gerald Friedland

in recognition of his contributions to organizing the IEEE International Conference on Semantic Computing (ICSC 2007). Friedland served as the Program Coordination Co-Chair for this conference, which took place September 17-19, 2007.

featured research, con't.

Continued from Page 1

malware. The new implementation exploits spam feeds (sources of substantial spam traffic) by carefully studying the traffic, luring malware, and executing it in a controlled fashion to study its behavior.

BOTS

While worms and viruses are the kinds of malware most familiar to the public, the spread of “bots” represents a more significant problem today. Bots are stealthier than worms, as they don’t inflict any obvious damage when they infect a computer. Running in the background, they use the infected machine for all kinds of malicious activities, including stealing credit card and other personal information, sending out spam email, or overloading a web server by sending multiple requests to the server from many infected machines at once (termed a denial-of-service attack). Some bots simply observe everything happening on a computer - every keystroke made by every user - and are able to stealthily inflict damage by (for example) adding items to an order while a user is logged into an e-commerce site.

Kreibich and the ICSI team are also working with scientists at UCSD led by Professor Stefan Savage to study botnets - networks of machines infected by bots. Spammers look for a topic of interest to get users to read an infected email. Early last year a huge storm in Europe provided them with a perfect subject line; the bot known as “Storm” spread very quickly and quietly without noticeably causing any problems, spreading to many (some experts estimate over a million) machines before anyone realized there was a problem. With such wide distribution, Storm could be quite damaging if it were used maliciously. Its scope and threat have generated a lot of concern, and this concern has produced media attention that is often unverified or even inaccurate. Kreibich and the UCSD team are studying Storm to combat it and make sure that the information regarding it is accurate and useful.

Security work at ICSI includes practical applications to prevent malware from infecting computers as well as the more theoretical study of malware to better understand how it works, and thus inform the security software of the future.

BRO

Bro, developed by Vern Paxson, is a network intrusion detection system currently in use at Lawrence Berkeley Laboratory (LBL) and UC Berkeley. It detects and logs hostile activity by filtering network traffic and analyzing relevant events against a set of rules to interpret and describe the network activity. Based on the information Bro detects, it can either create alerts of potentially troublesome activity, or execute programs to terminate malicious connections and

block traffic from hostile hosts. A particular advantage of Bro is that its more analytical features allow it to detect network behaviors manifested by novel forms of attacks.

Nicholas Weaver and Robin Sommer are both refining the prototype. Weaver has been

incorporating hardware acceleration and improving the rules by which Bro recognizes and analyzes connections. This polishing of Bro’s detection methods along with the hardware acceleration will improve Bro’s efficiency in detecting attacks.

To achieve high performance, Bro currently operates on a cluster of machines. Sommer’s work with Bro will enable it to run on a multicore parallel processor, which will decrease its latency of communication and keep it capable of continuing to scale its performance with future technology. He also works closely with the cybersecurity operations staff at LBL in their ongoing use of Bro and its associated technology.

VAST (VISIBILITY ACROSS SPACE AND TIME)

The VAST project is improving and standardize network activity logs to help determine how/why network attacks succeed. It also seeks to standardize the way logs are recorded so that logs from different networks can be compared to each other when similar problems have occurred. This will improve the ability of a network’s operators to detect problems such as intrusions. With improved and standardized logs, network administrators can study past network traffic and flag certain types of events to see if they have

occurred on different sites, what happened as a result, and even trigger a response if it happens in the future. Automated analysis of these logs allows problems to be tracked very quickly. VAST can be used in conjunction with a network intrusion system such as Bro to analyze and track the data seen by the intrusion detection system. This synthesis of ICSI activities improves Bro's ability to detect intrusions and determine when and how they happened.

TROUBLESHOOTING

While projects such as Bro and VAST depend on detailed logging and analysis to identify and prevent network intrusions, a new project at ICSI seeks to incorporate troubleshooting as an essential element in designing network architecture. The scope of Troubleshooting is not limited to one particular domain, but aims to be evolvable and applicable across architectures. In order to accomplish this, events found with Bro and VAST will be distilled and used in communication between system components. Even further, repositories of these logs and their distillations will be made available to parties not involved in a given transaction; this will facilitate cooperative trouble-shooting while still protecting user privacy and competitive information. ICSI's work will lay the groundwork for both a new graduate-level course in troubleshooting and provide material for undergraduate labs to prepare students for practical issues in networking.

Although the troubleshooting effort is a relatively new direction at ICSI, the team has already had success in this area. In collaboration with scientists at the University of Washington, ICSI's security team developed a method to detect insertions onto web pages by ISPs. They found that some ISPs are inserting ads and other modifications to the web pages seen by their viewers, and that some of these modifications include vulnerabilities. By creating more ways to determine when things like this are happening, and studying which are benign and which are malicious or cause vulnerabilities, scientists can improve overall security for users.

GRAND CHALLENGE

Weaver is also currently in the process of designing a security-related Grand Challenge competition. These competitions, such as robot races involving

cars that drive themselves, offer significant prizes for the most successful solution to a problem, and these prizes often attract people with wide-ranging skills and approaches to a challenge. Thus, Grand Challenge competitions can be an effective way to get lots of people working on a scientific obstacle and get interesting results quickly. Using NSF funding, Weaver is planning a conference, tentatively scheduled for summer 2008, where competitors would assume control of a server for 24 hours and defend it against attacks from the judges as well as from autonomous teams of hackers. The results of this competition can potentially inform other security efforts at ICSI by illustrating the efficacy of novel attack and defense methods.

INTERNET BLACK MARKET

Paxson, former ICSI intern Jason Franklin, and scientists at CMU and UCSD are fighting back against the Internet black market. Many illegal items, such as spamming services, toolkits for building fishing sites, and laundered money are available online. Criminal investigators need better tools to combat the spread of these markets. The researchers are looking at the problem of Internet attacks from a completely new perspective. By observing the commercial ecosystem online, the team is measuring and cataloging activities of users who profit from illegal online activities, and are searching for potential vulnerabilities of the market.

Vendors on the black market can have reputation ratings, not unlike those found on eBay. Potential countermeasures reported by the group use these ratings. The first technique involves slandering the users' reputations so that potential buyers are less able to find vendors of illegal goods that appear reliable. The second is to create a large number of false vendor identities, and using these false identities to legitimize each other. These "verified" vendors get business as if they were genuine and fail to deliver the goods and services. Techniques like these would create a lack of trust among potential buyers of black market services, so that they are less likely to feel comfortable buying products and services from sellers. Countermeasures such as these initial ideas could undercut the reliability of the existing market, turning it into a lemon market. The researchers are still learning how the black market functions and identifying ways in which it is vulnerable.

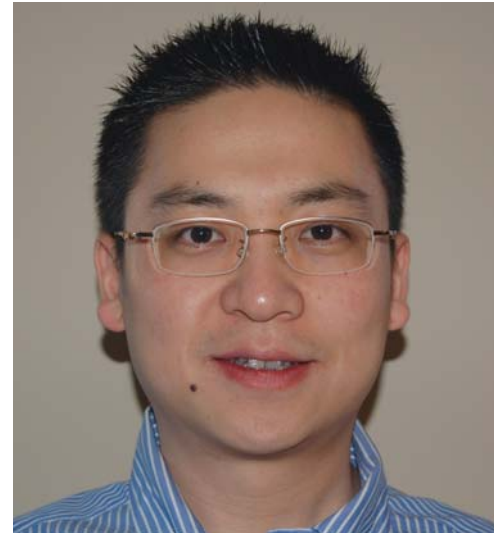
featured alum: weidong cui

Our featured alum for this issue is Weidong Cui, who developed the first version of ICSI's Honeyfarm technology (see front page story for more on the current Honeyfarm developments) while he was a research assistant with the Networking Group in 2005 and 2006.

Cui grew up in China, where he studied Electronic Engineering at Tsinghua University in Beijing, and received his BE and ME degrees. He then moved to California, where he completed his PhD at UC Berkeley in 2006. His thesis was focused on automatic malware detection, specifically, detection of malware by inferring the intent of the user or adversary. Most benign software running on personal computers is user driven, and the authors of different kinds of malware have distinct intent as to the type of damage they want to inflict.

While working on his PhD thesis at UC Berkeley, Cui was part of Vern Paxson's Internet Security team, working closely with both Paxson and Nicholas Weaver. "My working experience with Vern and Nick at ICSI was fantastic. Vern was my de facto co-advisor. I have learned tremendously from him: his high standard for research, exceptional diligence, efficient and effective communication, curiosity for new problems, and understanding and consideration for students. When I was working on the RolePlayer project, Vern gave me the freedom to pursue an approach he suspected at the beginning, which was very important for the success of the project. Nick was a nice and fun colleague to work with. He has generously helped me with my quals and research at ICSI. His reverse thinking always surprised me."

"My work at ICSI helped me improve and be ready for my current job in many perspectives...I went through the process of tackling a hard problem, from which I learned that it takes persistent efforts... and if I don't give up, the solution may be just steps away."



ICSI Alum Weidong Cui

At ICSI, Cui was responsible for much of the design, development, and deployment of the original Honeyfarm system, a project of the Collaborative Center for Internet Epidemiology and Defenses (CCIED), a cooperative effort between researchers at UCSD and ICSI. He also developed a novel system for automatically replaying network application dialog for use with the Honeyfarm.

After completing his PhD, Cui was hired by Microsoft Research, where he currently works on automatic reverse engineering of network protocols and file formats, as well as some security-related work on vulnerability signature generation, web security, hardware security, and botnet detection.

Cui says his time at ICSI was invaluable in preparing him for a private sector research job. "My work at ICSI helped me improve and be ready for my current job in many perspectives... I went through the process of tackling a hard problem (i.e., automatically replaying application level dialog based on two dialog samples), from which I learned that it takes persistent efforts to solve a hard problem, and if I don't give up, the solution may be just steps away."

new genetic study: lamp

Local Ancestry in admixed Populations

A new study by Dr. Eran Halperin of ICSI, along with his UCB and CMU colleagues, provides a means of pinpointing the ancestry of each position on an individual's genome. This information can be used to reveal complex ancestral history, which can then be used to study genetic diseases.

Large-scale genotyping of SNPs (single nucleotide polymorphisms, mutations that occurred once in history and were then passed on through heredity and became prevalent in a population) has been used extensively to identify markers that are associated with diseases. There are ~10 million SNPs in the human genome that differ between individuals.

The pattern of genetic differences between individuals varies across populations. In studies that involve more than one population, spurious associations between an SNP and a disease might be found due to the differences in these patterns and not due to the disease. In order to overcome this, methods exist to estimate the ancestry of each individual in a study, based on his or her genome. Until recently, accurate methods for finding the ancestry of individuals of recently admixed populations (such as African Americans, Latin Americans, and diverse urban populations such as San Francisco) did not exist. In these populations, each position in the genome could have been transmitted from a different ancestry.

The new method developed by Eran Halperin with Sriram Sankararaman (UC Berkeley), Srinath Sridhar (Carnegie Mellon University), and Gad Kimmel (ICSI and UC Berkeley) called Local Ancestry in admixed Populations (LAMP), which is published in [the February issue of American Journal of Human Genetics](#), is able to identify the ancestry of each position in the genome for individuals from admixed populations. The researchers performed extensive simulation studies to evaluate their method, and have shown that it is highly accurate and can precisely identify ancestral regions in the genome for typical cases of individuals that are a mix of two or three populations.

The main motivation for applying LAMP to genetic data is to avoid spurious results in disease association studies. However, LAMP will also be useful in improving studies of diseases that are especially



Eran Halperin

prevalent in specific populations. For instance, Multiple Sclerosis (MS) is much more common in Northern European populations than in others, and Alzheimer's disease is more common in African Americans than in Caucasians. The combination of this variation in prevalence, together with methods such as LAMP, can be used to study these diseases through studies known as admixture mapping. As an

LAMP uses whole-genome scans to pinpoint the ancestry of an individual in each position of the genome. This information can be used to reveal our ancestral history, which can then be used to study complex genetic diseases such as Multiple Sclerosis and Alzheimer's disease.

example, in MS one can genotype African-American cases and search for regions in the genome that are of Northern European ancestry. Since MS is more prevalent in Northern Europeans than in Africans, such genomic regions are suspected to be related to MS. The LAMP method accurately provides the ancestry of every single SNP in a person's genotype, enabling scientists to perform such admixture mapping studies more accurately than ever before.

vern paxson wins

grace murray hopper award



Vern Paxson

Vern Paxson, Senior Scientist with the Networking Group and a UCB Professor, is the recipient of the 2007 Grace Murray Hopper Award from ACM for outstanding young computer professional of the year. Paxson was selected for work he did on measuring Internet behavior. ACM issued a press release about Paxson's award on February 21st. Last year's winner, Dan Klein, is an ICSI Faculty Associate working with the Speech Group on machine translation.

ACM's press release follows:
(reprinted with permission)

The Association for Computing Machinery (ACM) has recognized Vern Paxson, Senior Scientist at the International Computer Science Institute (ICSI) Center for Internet Research in Berkeley, Calif., for his research on how to measure Internet behavior. His innovative techniques are used to assess new communications concepts, improve network performance, and prevent network intrusion. They provide both the research community and Internet operators with the tools

to improve the operation of this increasingly diverse, decentralized communications infrastructure. Paxson is also Associate Professor of Computer Science at the University of California, Berkeley, and a staff scientist at the Lawrence Berkeley National Laboratory. He will receive the 2007 Grace Murray Hopper Award from ACM for outstanding young computer professional of the year. The award carries a \$35,000 prize, and funding is provided by Google, Inc. Paxson's research on Internet measurement brought the scientific process to the measurement of the Internet's behavior and the conditions under which it operates, raising the practice of Internet measurement to a higher level. As a result, the research community is able to evaluate new ideas and technologies and identify problems and priorities that are needed for increased efficiency. In addition, Internet operators are able to alleviate traffic congestion, detect attacks, and improve communications reliability.

Through a series of highly influential papers, Paxson's findings revealed the mismatches between reality and the common assumptions made in analytical and simulation models. By combining the extensive collection of data from many locations with sophisticated statistical techniques, he provided a wealth of useful information about the nature of the Internet and ways to improve its operation.

Paxson was named an ACM Fellow in 2006. His 1996 research paper titled [End-to-end routing behavior](#)

[in the Internet](#) won the first "Test of Time" award given by ACM's Special Interest Group on Data Communication (SIGCOMM). The award, presented in 2006, is given to the most influential networking paper published 10-12 years before. His current research continues to focus on Internet measurement as well as network intrusion detection and large-scale Internet attacks.

In 2001, Paxson co-founded the Internet Measurement Conference and co-chaired the Program Committee of SIGCOMM 2002. He was an editorial board member of the IEEE/ACM Transactions on Networking from 2000-2004, and an organizer of the 2003 Workshop on Large-scale Internet Attacks at Rutgers University's Center for Discrete Mathematics and Computer Science (DIMACS). From 2001-2005, he chaired the Internet Research Task Force (IRTF), which promotes research critical to the evolution of the future Internet.

Paxson received his M.S. and Ph.D. degrees from the University of California, Berkeley.

The Grace Murray Hopper Award honors the outstanding young computer professional of the year, selected on the basis of a single recent major technical or service contribution. The candidate must have been 35 years of age or less at the time the qualifying contribution was made. ACM will present the 2007 Grace Murray Hopper Award to Paxson at the annual ACM Awards Banquet on June 21, 2008, in San Francisco, CA.

visiting scholars

Since its inception, ICSI has had a strong international program consisting primarily of ties with specific countries. Current formal agreements exist with Finland, Germany, Spain, and Switzerland.

FROM CHINA

Bin Dai (Networking)

FROM EUROPEAN UNION (AMIDA)

Joan Isaac Biel (Speech)
Thomas Kleinbauer (AI-FrameNet)
Jachym Kolar (Speech)
Rosemary Orr (Speech)
Korbinian Riedhammer (Speech)
Beatriz Trueba (Speech)
David VanLeeuwen (Speech)
Oriol Vinyals (Speech)

FROM FINLAND

Ari-Veikko Anttiroiko (Campus Affiliation)
Jyri Kivinen (Algorithms/Campus Affiliation)
Teemu Koponen (Networking)
Tommi Lampikoski (Campus Affiliation)
Annukka Näyhä (Campus Affiliation)
Teemu Roos (Campus Affiliation)
Ville-Pekka Seppä (Campus Affiliation)
Pekka Valkama (Campus Affiliation)

FROM GERMANY

Gerald Friedland (Speech)
Martin Gairing (Algorithms)
Martin Hilpert (AI-FrameNet)
Christian Kreibich (Networking)
Birte Lönneker-Rodman (AI-FrameNet)
Andreas Maletti (AI)
Christian Müller (Speech)
Felix Salfner (Campus Affiliation)
Vladimir Stantchev (Networking)

FROM SPAIN

Carlos Subirats (AI-FrameNet)

FROM SWITZERLAND (IM2)

Neha Garg (Speech)
Bao-Lan Huynh (Speech)
Adish Singla (Speech)
Kamand Kamangar (Speech)



Bin Dai



Annukka Näyhä



Felix Salfner

In addition, we often have visitors associated with specific research and projects.

AI

Alberto Amengual
Artjom Klein

ALGORITHMS

Lucia Conde
Ron Shamir

EXTREME ARCHITECTURE

Chris Batten
Rose Liu
Heidi Pan

FRAMENET

Santi Caballe
Seiko Fujii

NETWORKING

Andrew Che
Leo Juan
Po-Ching Lin
Li Tang
Shutang Yang

NETWORKING INTERNS

Matthias Vallentin

SPEECH

Michael Feld
Umit Guz
Oded Ghitza

publications listing

- M. ALLMAN, K. CHRISTENSEN, B. NORDMAN, AND V. PAXSON. ENABLING AN ENERGY-EFFICIENT FUTURE INTERNET THROUGH SELECTIVELY CONNECTED END SYSTEMS. PROCEEDINGS OF HOTNETS 2007.
- M. ALLMAN AND V. PAXSON. ISSUES AND ETIQUETTE CONCERNING USE OF SHARED MEASUREMENT DATA. PROCEEDINGS OF ACM INTERNET MEASUREMENT CONFERENCE, OCTOBER 2007.
- M. ALLMAN, V. PAXSON, AND J. TERRELL. A BRIEF HISTORY OF SCANNING. PROCEEDINGS OF ACM INTERNET MEASUREMENT CONFERENCE, OCTOBER 2007.
- K. ASANOVIĆ. TRANSACTORS FOR PARALLEL HARDWARE AND SOFTWARE CO-DESIGN. IEEE INTERNATIONAL HIGH LEVEL DESIGN VALIDATION AND TEST WORKSHOP 2007 (HLDVT-2007), (INVITED PAPER), IRVINE, CA, NOVEMBER 2007.
- K. ASANOVIĆ AND E. WITCHEL. SYSTEM AND TECHNIQUE FOR FINE-GRAINED COMPUTER MEMORY PROTECTION. US PATENT 7,287,140, GRANTED OCTOBER 2007.
- L. BARRETT, J. FELDMAN, AND L. MAC DERMED. A (SOMEWHAT) NEW SOLUTION TO THE BINDING PROBLEM. NEURAL COMPUTATION, 2008 (PREVIOUSLY PUBLISHED AS ICSI TECHNICAL REPORT IN 2006)
- D.C. CHU, L. POPA, A. TAVAKOLI, J.M. HELLERSTEIN, P. LEVIS, S. SHENKER, AND I. STOICA. THE DESIGN AND IMPLEMENTATION OF A DECLARATIVE SENSOR NETWORK SYSTEM. PROCEEDINGS OF 5TH INTERNATIONAL CONFERENCE ON EMBEDDED NETWORKED SENSOR SYSTEMS (SENSYS '07), NEW YORK, PP. 175-188, 2007.
- M. CREUTZ, T. HIRSIMÄKI, M. KURIMO, A. PUURULA, J. PYLKKÖNEN, V. SIVOLA, M. VARJOKALLIO, E. ARISOY, M. SARAÇLAR, AND A. STOLCKE. MORPH-BASED SPEECH RECOGNITION AND MODELING OF OUT-OF-VOCABULARY WORDS ACROSS LANGUAGES. ACM TRANSACTIONS ON SPEECH AND LANGUAGE PROCESSING 5(1), ARTICLE 3, 29 PAGES, 2007.
- A. FARIA AND N. MORGAN. WHEN A MISMATCH CAN BE GOOD: LARGE VOCABULARY SPEECH RECOGNITION TRAINED WITH IDEALIZED TANDEM FEATURES. PROCEEDINGS OF THE ACM SYMPOSIUM ON APPLIED COMPUTING, FORTALEZA, BRAZIL, MARCH 2008.
- J. FEIGENBAUM, M. SCHAPIRA, AND S. SHENKER. DISTRIBUTED ALGORITHMIC MECHANISM DESIGN. BOOK CHAPTER IN ALGORITHMIC GAME THEORY. NOAM NISAN, Ed., CAMBRIDGE UNIVERSITY PRESS, 2007.
- C.J. FILLMORE. VALENCY ISSUES IN FRAMENet. IN THOMAS HERBST & KATRIN GÖTZ-VOTTELER, EdS: VALENCY; THEORETICAL, DESCRIPTIVE AND COGNITIVE ISSUES, MOUTON DE GRUYTER, THE HAGUE, PP. 128-160, 2007.
- J. FRANKLIN, V. PAXSON, A. PERRIG, AND S. SAVAGE. AN INQUIRY INTO THE NATURE AND CAUSES OF THE WEALTH OF INTERNET MISCREANTS. PROCEEDINGS OF ACM CCS, OCTOBER 2007.
- G. FRIEDLAND AND R. ROJAS. ANTHROPOCENTRIC VIDEO SEGMENTATION FOR LECTURE WEBCASTS. EURASIP JOURNAL ON IMAGE AND VIDEO PROCESSING, HINDAWI PUBLISHING, 2008.
- J. GONZALEZ, V. PAXSON, AND N. WEAVER. SHUNTING: A HARDWARE/SOFTWARE ARCHITECTURE FOR FLEXIBLE, HIGH-PERFORMANCE NETWORK INTRUSION PREVENTION. PROCEEDINGS OF ACM CCS, OCTOBER 2007.
- D. HAKKANI-TÜR, H. JI, AND R. GRISHAM. USING INFORMATION EXTRACTION TO IMPROVE CROSS-LINGUAL DOCUMENT RETRIEVAL. PROCEEDINGS OF RECENT ADVANCES IN NATURAL LANGUAGE PROCESSING (RANLP) WORKSHOP ON MULTILINGUAL INFORMATION EXTRACTION AND SUMMARIZATION, BULGARIA, 2007.
- M. HILPERT AND S. GRIES. THE IDENTIFICATION OF STAGES IN DIACHRONIC DATA: VARIABILITY-BASED NEIGHBOR CLUSTERING. CORPORA, ISSUE 3, VOLUME 1, 2008.
- M. HILPERT. THE ENGLISH COMPARATIVE - LANGUAGE STRUCTURE AND LANGUAGE USE. ENGLISH LANGUAGE AND LINGUISTICS, 2008.
- M. HOLMBERG, D. GELBART, AND W. HEMMERT. SPEECH ENCODING IN A MODEL OF PERIPHERAL AUDITORY PROCESSING: QUANTITATIVE ASSESSMENT BY MEANS OF AUTOMATIC SPEECH RECOGNITION. SPEECH COMMUNICATION VOLUME 49, ISSUE 12, DECEMBER 2007.
- Y. HUANG, O. VINYALS, G. FRIEDLAND, C. MÜLLER, N. MIRGHAFORI, AND C. WOOTERS. A FAST-MATCH APPROACH FOR ROBUST, FASTER THAN REAL-TIME SPEAKER DIARIZATION. PROCEEDINGS OF IEEE 10TH BIENNIAL WORKSHOP ON AUTOMATIC SPEECH RECOGNITION AND UNDERSTANDING (ASRU), KYOTO, JAPAN, 2007.
- M-Y. HWANG, G. PENG, W. WANG, A. FARIA, AND A. HEIDEL. BUILDING A HIGHLY ACCURATE MANDARIN SPEECH RECOGNIZER. PROCEEDINGS OF IEEE 10TH BIENNIAL WORKSHOP ON AUTOMATIC SPEECH RECOGNITION AND UNDERSTANDING (ASRU), KYOTO, JAPAN, 2007.
- S. KIM, R. FONSECA, P. DUTTA, A. TAVAKOLI, D. E. CULLER, P. LEVIS, S. SHENKER, AND I. STOICA. FLUSH: A RELIABLE BULK TRANSPORT PROTOCOL FOR MULTIHOP WIRELESS NETWORKS. PROCEEDINGS OF 5TH INTERNATIONAL CONFERENCE ON EMBEDDED NETWORKED SENSOR SYSTEMS (SENSYS '07), NEW YORK, PP. 175-188, 2007.
- G. KIMMEL, M. JORDAN, E. HALPERIN, R. SHAMIR, AND R.M. KARP. A RANDOMIZATION TEST FOR CONTROLLING POPULATION STRATIFICATION IN WHOLE-GENOME ASSOCIATION STUDIES. THE AMERICAN JOURNAL OF HUMAN GENETICS, v. 81, n. 5, PP. 895-905, Nov. 2007.
- B. KIRKPATRICK, C.S. ARMENDARIZ, AND R.M. KARP. HAPLOPOOL: IMPROVING HAPLOTYPE FREQUENCY ESTIMATION THROUGH DNA POOLS AND PHYLOGENETIC MODELING. BIOINFORMATICS, SEPTEMBER 2007.
- J. KIVINEN, E. SUDDERTH, AND M. JORDAN. IMAGE DENOISING WITH NONPARAMETRIC HIDDEN MARKOV TREES. PROCEEDINGS OF ICIP 2007.
- J. KIVINEN, E. SUDDERTH, AND M. JORDAN. LEARNING MULTISCALE REPRESENTATIONS OF NATURAL SCENES USING DIRICHLET PROCESSES. PROCEEDINGS OF INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV), OCTOBER 2007.
- J.W. LEE, M. KING, AND K. ASANOVIĆ. CONTINUAL HASHING FOR EFFICIENT FINE-GRAIN STATE INCONSISTENCY CHECKING. IEEE INTERNATIONAL CONFERENCE ON COMPUTER DESIGN (ICCD-2007), LAKE TAHOE, CA, OCTOBER 2007.

M. LEVIT, D. HAKKANI-TUR, AND G. TUR. INTEGRATING SEVERAL ANNOTATION LAYERS FOR STATISTICAL INFORMATION DISTILLATION. PROCEEDINGS OF IEEE 10TH BIENNIAL WORKSHOP ON AUTOMATIC SPEECH RECOGNITION AND UNDERSTANDING (ASRU), KYOTO, JAPAN, 2007.

B. LÖNNEKER-RODMAN. BEYOND SYNTACTIC VALENCE: FRAME-NET MARKUP OF EXAMPLE SENTENCES IN A SLOVENIAN-GERMAN ONLINE DICTIONARY. IN JANA LEVICKÁ AND RADOVAN GARABIK (EDS.): SLOVKO 2007. COMPUTER TREATMENT OF SLAVIC AND EAST EUROPEAN LANGUAGES. PROCEEDINGS OF THE FOURTH INTERNATIONAL SEMINAR. BRATISLAVA, SLOVAKIA, TRIBUN, BRATISLAVA 152-164, OCTOBER 2007.

B. LÖNNEKER-RODMAN, C. BAKER, AND J. HONG. THE NEW FRAME-NET DESKTOP: A USAGE SCENARIO FOR SLOVENIAN. IN PROCEEDINGS OF ICGL 2008, THE FIRST INTERNATIONAL CONFERENCE ON GLOBAL INTEROPERABILITY FOR LANGUAGE RESOURCES, HONG KONG, 147-154, JANUARY 2008.

A. LÜNING, G. FRIEDLAND, L. KNIPPING, AND R. ROJAS. VISUALIZING LARGE-SCREEN ELECTRONIC CHALKBOARD CONTENT ON HANDHELD DEVICES. PROCEEDINGS OF THE 2ND IEEE INTERNATIONAL WORKSHOP ON MULTIMEDIA TECHNOLOGIES FOR E-LEARNING AT 9TH IEEE SYMPOSIUM ON MULTIMEDIA, PP. 369-375, TAICHUNG, TAIWAN (R.O.C), DECEMBER 2007.

J. MAKIN AND S. NARAYANAN. HYBRID-SYSTEM MODELING OF HUMAN BLOOD CLOTTING. ICSI TECHNICAL REPORT TR-08-002, BERKELEY, CA, FEBRUARY 2008.

A. MALETTI. MYHILL NERODE THEOREM FOR RECOGNIZABLE TREE SERIES --- REVISITED. PROCEEDINGS OF 8TH INTERNATIONAL CONFERENCE ON LATIN AMERICAN THEORETICAL INFORMATICS, 2008.

A. MALETTI. MINIMIZING DETERMINISTIC WEIGHTED TREE AUTOMATA. PROCEEDINGS OF 2ND INTERNATIONAL CONFERENCE ON LANGUAGE AND AUTOMATA: THEORY AND APPLICATIONS, 2008.

J. PAL. EXAMINING E-LITERACY USING TELECENTERS AS PUBLIC SPENDING. INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGIES AND DEVELOPMENT (ICTD2007), IEEE CONFERENCE PROCEEDINGS, BANGALORE, INDIA, DECEMBER 2007.

J. PAL, M. LAKSHMANAN, AND K. TOYAMA. 'MY CHILD WILL BE RESPECTED': PARENTAL PERSPECTIVES ON COMPUTERS IN RURAL INDIA. INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGIES AND DEVELOPMENT (ICTD2007), IEEE CONFERENCE PROCEEDINGS, BANGALORE, INDIA, DECEMBER 2007.

R. PATRA, J. PAL, S. NEDEVSCHI, M. PLAUCHÉ, AND U.S. PAWAR. USAGE MODELS OF CLASSROOM COMPUTING IN DEVELOPING REGIONS. INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGIES AND DEVELOPMENT (ICTD2007), IEEE CONFERENCE PROCEEDINGS, BANGALORE, INDIA, DECEMBER 2007.

S. SANKARARAMAN, S. SRIDHAR, G. KIMMEL, AND E. HALPERIN. LAMP: LOCAL ANCESTRY IN ADMIXED POPULATIONS. THE AMERICAN JOURNAL OF HUMAN GENETICS, VOLUME 82, ISSUE 2, 290-303, FEBRUARY 2008.

D. THAW, J. FELDMAN, AND J. LI. CoPE: DEMOCRATIC CSCW IN SUPPORT OF E-LEARNING. PROCEEDINGS OF IEEE CISIS, BARCELONA, SPAIN, MARCH 2008.

M. VALLENTIN, R. SOMMER, J. LEE, C. LERES, V. PAXSON, AND B. TIERNEY. THE NIDS CLUSTER: SCALABLE, STATEFUL NETWORK INTRUSION DETECTION ON COMMODITY HARDWARE. PROCEEDINGS OF RAID 2007.

O. VINYALS AND G. FRIEDLAND. LIVE SPEAKER IDENTIFICATION IN MEETINGS: "WHO IS SPEAKING NOW?" ICSI TECHNICAL REPORT TR-08-001, BERKELEY, CA, JANUARY 2008.



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 four 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; and natural speech processing.

BOARD OF TRUSTEES

Javier Aracil
Hervé Bourlard
Beth Burnside
Adele Goldberg
Greg Heinzinger
Clifford Higgerson
Richard Karp
Nelson Morgan
David Nagel
Prabhakar Raghavan
Stuart Russell
Jouko Salo
Shankar Sastry
Scott Shenker
David Tennenhouse
Wolfgang Wahlster

The ICSI Gazette is published twice a year by the International Computer Science Institute (ICSI). All material in The ICSI Gazette is copyrighted by ICSI and should not be used without explicit prior written permission.

EDITOR-IN-CHIEF

Nelson Morgan

SENIOR EDITOR

Maria Quintana

DESIGN

envision ID

LAYOUT

Leah Hitchcock

COPYWRITER

Leah Hitchcock

GUEST COPYWRITER

Jacob Wolkenhauer

PHOTOGRAPHY

Diane Starr

INTERNATIONAL
COMPUTER
SCIENCE
INSTITUTE

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