the newsletter of the International Computer Science Institute

The ICSI GAZETTE

volume eleven | issue one | september 2012

featured icsi research: speech

Speech recognition has recently become a popular topic, with Apple's Siri and other voice assistants making frequent appearances in the news. But speech technology's recent surge in popularity isn't related to any major breakthroughs in speech technology. Rather, advances in speech recognition have been incremental, and, according to Speech Group researchers, there's still plenty of work to do before technology can understand human speech as well as humans.

"Machine intelligence fails at all sorts of things that humans don't," said Nelson Morgan, who has led the group since its formation in 1988.

ICSI has had its share of success in the 24 years that researchers here have been working on speech recognition. What is now the Speech Group began as the Realization Group, reflecting its focus on building, or realizing, machines powerful enough to process the algorithms used in research. But even in the early years, researchers were interested in problems related to speech processing.

Focus on Features

In a speech recognition system, audio is segmented into sequential chunks of speech, and features are extracted from each chunk. Ideally, these features are chosen to be good for distinguishing between different speech sounds (for instance, between "ba" and "ga"). On the other hand, they should be insensitive to noise and other non-speech effects, though this can be difficult.

An acoustic model, which in most systems is a Hidden Markov Model (HMM), then determines the probability that each audio chunk is a particular sound based on its associated features. Independently, a language model determines the probability of a sequence of words without considering the acoustics. For example, the phrase "I want ice cream" occurs more commonly than "I won Ty's ream"; a language model would favor the first phrase, even if the acoustic model favored the second.

Morgan says speech research of the 1980s and 1990s may have suffered from too great an emphasis on refining these probabilistic models. "How you use features and what features you use are critically important to all sorts of tasks," he said.

In 1990, ICSI hosted a workshop dedicated to front-end speech processing. At the workshop, Jordan Cohen (later to become a frequent ICSI collaborator), presented the "Problem of the Inverse E": if you build a system to filter out the spectrum of the sound "E" from a speech data set, a human listener can still hear the "E's".

"We realized that the perception of each speech sound category from continuous speech couldn't possibly be due to the gross spectrum of the sound in the small chunks that are used in speech recognition systems," Morgan said. "This was a heretical view at the time."

Morgan worked with Hynek Hermansky, then a researcher at US West, to develop a new way of analyzing audio, called Relative Spectral Processing or RASTA. As opposed to standard approaches, this method was more sensitive to spectral changes over time and less sensitive to the spectrum itself. This kind of processing helps systems handle the sometimes drastic differences in spectrum between the data recorded training the models and the data used for testing them.

For example, most speech systems at the time had difficulty recognizing audio recorded on different microphones from those used to record its training data. With RASTA, what's important is the change from one moment to the next, not the absolute audio spectrum at any given point. This means that differences between microphones become less important in the speech recognition process. Morgan said the early work

continued on page 4

photo by Leide Porcu

reflections by Roberto Pieraccini, Director

W. Brian Arthur, in his book *The Nature of Technology: What It Is and How It Evolves*, describes the evolution of technology as a combinatorial process. Each new technology consists in a combination of existing technologies that "beget further technologies." Moreover, each technology springs from the harnessing of one or more physical, behavioral, mathematical, or logical principles that constitute its foundation. Innovation proceeds either through the establishment of new principles—which is typically, but not only, the domain of science—or through new combinations of existing technologies.

Not all technologies are alike, however; sometimes a single new technology, anenabler, generates a myriad of new possibilities that lead to the creation of new industries and new economies, and in some very rare occasions contribute to the definition of a new era of our civilization. Such were the steam engine, the digital computer, and the Internet.

We may find ourselves wondering what the next enabler will be. Of course no one knows for sure, and any attempt to make a prediction will most likely be wrong. But researchers have a special role in our technological future. They cannot predict what the future will be but, to paraphrase Alan Kay's famous quote, they can attempt at creating it. Well, looking at the trends of current research in information technology, we can definitely see that the attempt to create a new future based on automatically deriving higher levels of data understanding is today one of the most challenging endeavors researchers are embarked on. Let me be more specific.

Our era is characterized by an unprecedented amount of information. It is no surprise that a significant amount of technological research today is devoted to the creation, management, and understanding, by computers, of the wealth of data – text, images, and sounds around us. But it is the understanding, which is the most challenging among those problems and the farthest from a satisfactory solution. A large portion of the research community aims to devise algorithms and systems to automatically extract the meaning, the semantics, from raw data and signals. In fact, a lot of the research carried out at ICSI, as at many other research centers, can be described as "looking for meaning in raw data." Research on natural language and visual and multimedia signals is diving into the problem of deeper understanding. Beyond the mere (so to speak) recognition of words and entities in language and objects in images, we are now trying to get to the deeper content, such as metaphors and visual concepts. And beyond that, the research in network security, for instance, is trying to assign meanings to the patterns found in streams of communication among computers in order to detect possible intrusions and attacks, while theoretical computer scientists are trying to find meaning in DNA strings and in networks of brain activity.

However, we are not quite there yet. Think, for instance, about the promises made by the vision of the semantification of the whole Web. The vast majority of the Web comprises unstructured raw data: text, images, audio and videos. Tim Berners Lee was the first to envision a semantic web, and many have been working toward that dream, with limited degrees of success. Even though many agree on ways to encode a semantic web, and Google's knowledge graph is an example of one of the most advanced large-scale attempts to structure factual knowledge and make it readily available to everyone, a full semantic Web is not there yet. The knowledge graph starts from existing structured knowledge, for instance the facts about Albert Einstein life, and connects that structured knowledge to Web searches. The knowledge graph includes millions of entries, which is an infinitesimally small number compared to the vast universe of the Web. Are you and your papers on the knowledge graph? Are all recent world facts, blog entries, and opinions on the European financial crisis in the world graph? Maybe they will be, but the question of coverage and that of keeping the information fresh and updated is yet to be solved. And an even more serious issue: the Web is not just text. For instance, the amount of video on the Web is growing at a mindboggling rate. Some of the recent published statistics about YouTube estimate that 72 hours of video are uploaded every minute, over 4 billions hours are watched every month, and in 2011 alone, YouTube had more than 1 trillion views, the equivalent of over 140 views for every person on Earth.

It is true that we have ways to encode semantics into Web pages, as seen in the work of W3C; semantic representation languages, like OWL are widely used today. But with the Web growing at dizzying speed any attempt to manually annotate every unstructured element—text or video—with its meaning or something closely related to it, is bound to fail. If we want to fulfill the dream of a fully semantic Web, we need methods for automatically understanding text, images, and videos, including speech, music, and general audio.

The enabling potential of a fully semantic Web is huge. A fully semantic Web will change the concept of searchingthe Web into that of asking questions of it and getting answers, not just, as is sometimes possible today, in restricted domains, but everywhere, about any topic, no matter the size, or popularity, or language. It will help transform mere facts into structured data and actionable information, not just about Albert Einstein and other famous people, but also about you, the grocery store at the corner, and what your friends write on their blogs. It will complete the process of moving from raw data—the gazillions of pages of online text and video—toward higher abstractions of knowledge, understanding, and wisdom. And that's not all. Semantification of the whole Web will enable the flourishing of new areas of the information industry that are not possible today, or that are possible with a lot of handcrafting and ad-hoc solutions in limited domains and hardly scalable to other domains, languages, and media. It will allow us to interact with the Web as we do with humans, asking it questions in the same way we ask human experts questions. It will allow us to automatically compare sources of information for accuracy and truthfulness, even if they are in different languages.

photo by Roberto Pieraccini

However, the true full semantification of the Web is not just an ambitious dream, but a necessity. We may reach a point, in the not too distant future, when the Web will be so large that current search methods will no longer be the most effective way to find the information we need in the format we need it. How do we sort through an increasingly vast number of documents, and not just text, on a particular topic, with many independent and conflicting opinions, comprising true and false statements and different points of view? We already have to use specialized search engines to find the cheapest flight among independent travel sites and aggregators and to find the most economical reseller of a particular item. That's possible today because of the underlying structured form of the commercial sites. In a way, they are already semantified. But think, for a minute, of doing the same thing with unstructured information: raw text, audio, images, and video. Think about searching for documents that "have the same meaning" of another document, regardless of their language, form, and wording.

I don't know if we will see a full semantification of the Web in our lifetimes. I don't know if that's even possible, or whether it is that great enabler we dream of. But one thing is certain: research is clearly moving toward a deeper understanding of the information around us, and if it is successful, we will be able to experience a higher level of social, economic and political influence of the technologies in our lives. The full semantification of the Web, whenever it will happen, will be a game changer of enormous proportions, an enabler of industries and services which will impact all aspects of our lives. We are working on it.

featured research: continued

continued from page 1

on RASTA features, as well as much more recent successes, stressed the importance of front-end processing. Speaking of the problem of training/test spectral mismatch, he said “We woke people up to the fact that this was a problem. We weren’t the first people to suggest that, but we were the first ones to talk about it so loudly.”

RASTA was used in ICSI’s Berkeley Restaurant Project (BeRP), a spoken dialog system that gave restaurant recommendations. The system was unusual in that both the system and its users could initiate questions, and the system could continue a conversation even when users did not respond directly to its questions. BeRP incorporated a speech recognizer, a natural language backend that parsed words and produced database queries, a restaurant database, and accent detection and modeling algorithms that helped the system understand foreign accents and nonstandard pronunciations. The system was developed by Morgan, postdoctoral fellows Gary Tajchmann and Dan Jurafsky, and graduate students Chuck Wooters and Jonathan Segal.

Wooters, who was Morgan’s first graduate student and who recently returned to ICSI as a senior researcher, said the system had a tight integration between natural language understanding and speech recognition. “You didn’t think of the speech recognizer as static,” he said. “It was more of a living system.”

Learning from Biological Systems

RASTA is an example of technology emulating human systems, a theme throughout much of the speech recognition work at ICSI. “It’s really important to pay attention to what mechanisms we can discover from biological systems,” Morgan said.

His doctoral thesis work was on digitally reproducing some of the effects of room acoustics on speech and music. His approach relied on aspects of human perception. When a sound is made inside a room, a human listener first hears the sound that travels to his ears directly, and then hears the reverberations as the sound bounces off different surfaces. An analysis in the 1960s by Leo Beranek, cofounder of BBN, had shown that the early sounds were special: concert halls preferred by conductors were similar in that the time between the first sound and the first reverberation was about the same. Morgan said he realized, “Maybe there’s something critically important about those first sounds.” He built hardware and software to reproduce the early sounds in detail, while using a coarser method to approximate the later reverberation. The system performed well: in listening tests, study participants highly correlated the reproduced sounds with the correct room characteristics. The work incorporated ideas from psychoacoustics, which studies how audio stimuli affect perceptual processes.

The Speech Group is also interested in the effect that physiology has on the perception of audio. Research has shown that certain parts of the auditory system, of both humans and animals, are more attuned to certain aspects of audio. “The physiology gives a clue to some things that were harder to notice with just perception,” Morgan said. “It’s a potential source of inspiration for the kinds of things that we want to achieve.”

Early work on machine learning was greatly inspired by models of neurons, and artificial neural networks have been used for parts of some speech recognition systems for decades. Morgan has worked with neural networks since his days as a researcher at National Semiconductor, where he used neural networks in a speech analysis system. Then, in the 1990s, Morgan collaborated with Hervé Bourlard, now the director of IDIAP, on a hybrid approach that used neural networks statistically with HMMs. HMMs give the probability that a piece of sound is a particular word (or part of a word or a sentence). They are used in almost all current speech recognizers.

HMMs require a set of acoustic probabilities (that is, how likely it is that a chunk of sound corresponds to a particular speech sound like “ba”). In Bourlard and Morgan’s hybrid system, those probabilities are determined by an artificial neural network. Bourlard and Morgan’s paper on the approach, summarizing their joint work over the previous 7 years, won an *IEEE Signal Processing Magazine* best paper award from the Signal Processing Society in 1996, and their work together inspired other research directions throughout the 1990s. The hybrid approach is now experiencing a comeback with the growing popularity of work on deep learning.

The ICSI Meeting Corpus

By the mid 1990s, the Speech Group was looking for more difficult problems. Morgan said, "We were mostly looking at robustness in some sense – why are speech recognition systems breaking down? How do you make them less sensitive?"

A student suggested that Morgan, who was on his way to a meeting in Europe, keep track of times when speech recognition would have been useful for a handheld device, for instance to do calendar functions (such as now can be done with Siri). When he returned, Morgan realized he needed, not a personal electronic assistant, but some easy way of recording and retrieving notes from the meeting.

"All of the sudden it struck me: that's the application that would appeal to me. You want to be able to have access to information from some extended meeting or meetings by querying for it," he said. Perhaps more importantly for Morgan, such an application would drive research in many areas.

From this idea emerged the ICSI Meeting Corpus, a collection of recorded audio from meetings held at the Institute, along with transcriptions to aid in training speech recognition systems. At the time, it was the largest corpus of transcribed meetings available.

It was important that these recordings were of spontaneous speech. It included laughter, speech from multiple people talking at the same time, and vocalized pauses – "ums" and so forth. It also included speech recorded far from the speaker. These elements made for interesting problems in speech recognition, which the team set about solving.

Fundamental Problems

The main focus of Morgan's work has always been to solve fundamental problems rather than to push for incremental improvements in speech technology. For example, Speech Group researchers are currently developing ways to build speech recognizers for languages that don't have much training data. Morgan said this will force the team to significantly alter speech recognition methods.

"There's a lot to be gained in using lots of data," he said. "But it can mask how dumb your model is."

Another problem facing speech recognizers is how to handle speech with lots of background noise. Humans are generally pretty good at distinguishing noise and speech; machines are not. "The best-functioning systems try to get around this by having people speak directly into microphones," Morgan said. But that technique, of course, doesn't address the fundamental problems or cover many practical situations.

One problem is that speech recognition relies on algorithms that have been used since the 1960s. Speech Group researcher Steven Wegmann said that HMMs rely on assumptions that are "really strong and really wrong."

Morgan said in a recent interview on Marketplace Tech that improvements to speech technology come slowly because the algorithms used for speech recognition – HMMs – have not changed since the 1960s. "The major source of improvements has been the speed-up improvements in computers," he said. "But it's still the same basic fundamental algorithms."

Now, Wegmann and his colleagues at ICSI are trying to figure out what's wrong with HMMs. They will do this by simulating data and statistically analyzing errors. The project also includes a survey of experts in speech recognition to get their opinions on where the technology is failing, what has been tried that didn't work, and what still looks promising.

featured researcher:
chuck wooters

This summer, ICSI welcomed back on of its first employees: Chuck Wooters. This is Chuck's third time at ICSI.

Chuck grew up in the San Gabriel Mountains in southern California and entered UC Berkeley as an undergraduate. He planned to get his bachelor's degree in philosophy and then go to law school, but after taking a course in philosophical linguistics, he decided to switch majors to linguistics. His plans changed again when he became interested in computer science. He decided he would get his law degree, work as an attorney for a while, and pursue computer science later. But a linguistics professor, William Wang, pointed out that the work load of being an attorney would prevent him from developing his interest in computer science. Wang said, "You'll never do it if you get a law degree."

Chuck decided to sign up for the GREs, which were only a couple weeks away, and later entered the graduate program in linguistics at UC Berkeley. He received his master's degree in 1988 but decided didn't want a PhD in linguistics because he wasn't sure he'd be able to find a job working with computers. But he couldn't switch to computer science since he hadn't taken the math courses required by the department. In the end, he decided to design his own PhD program.

"Really what I was interested in was speech recognition," he said.

So he created a program and pulled together a committee of professors from the Computer Science, Linguistics, and Psychology Departments. He became the first – and possibly the only – student to receive a doctorate in "speech recognition."

He was also the first graduate student supervised by Nelson Morgan, the leader of the Realization Group (which would later become the Speech Group). While working on his master's during a semester abroad in Taiwan, Chuck had built a neural network classifier to distinguish B, D, and G sounds. This impressed Morgan, who used neural networks extensively in his speech recognition research.

Chuck joined ICSI in 1989 and worked here through the granting of his doctorate in 1993. "I didn't need the support of a department because I was here at ICSI," he said. "I felt like this was my department."

Morgan said, "Chuck's skills spanned a range from computer programming to linguistics, and he picked up all the necessary speech engineering tools he needed over the next few years."

Among his skills, Morgan said, is his ability to come up with bad puns. "This was part of his charm," Morgan said. "He was really quite wonderful to have around."

While at ICSI, he worked on the Berkeley Restaurant Project (BeRP), a spoken dialogue system that made restaurant recommendations based on criteria that users spoke into it. Chuck worked with Dan Jurafsky, at the time an ICSI postdoc who is now a professor at Stanford, to integrate a natural language understanding parser into the speech recognition system.

Chuck went to the U.S. Department of Defense, where he worked on applied speech recognition research. He then worked at a start-up that was building a voice-activated robot arm used in medical surgeries. It's the only speech recognition system approved by the Food and Drug Administration.

He also worked at the Department of Defense a second time and then BBN Technologies before returning to ICSI, where, among other things, he worked on the initial software for the FrameNet database.

Then, in mid-2007, he accepted a position at Next IT Corp, which builds and maintains virtual assistants on corporate Web sites. Clients include major airlines and the U.S. Army. To use the assistants, users type questions into a dialog box, and a natural language understanding (NLU) system analyzes the questions and provides a response. His first job at Next IT was to build a bridge between the NLU system and a speech recognition system so that users could speak their questions. But the company soon realized it could never compete with commercial speech recognition powerhouses like Google, AT&T, and Nuance, so it decided to focus exclusively on NLU. Among other things, Chuck worked on an algorithm to do parallel clustering of large quantities of text.

In the summer of 2009, he was given leave to attend a ten-week workshop at the Human Language Technology Center of Excellence at Johns Hopkins. It was the first iteration of the Summer Camp for Advanced Language Exploration, which has been held every year since.

"I now have a perspective that includes for-profit, government, and academic research." he said. While he has enjoyed his time in applied research, "I love the creative atmosphere and freedom provided by academic research."

In July 2012, he returned again to ICSI. "It's like coming home," he said. "I grew up at ICSI." Morgan says Chuck was his "baby," and in his office, Morgan still displays a pair of tennis shoes that he had bronzed when Chuck graduated.

> "I didn't need the support of a department because I was here at ICSI. I felt ICSI was my department."

Right now, Chuck is relearning the tools used for speech recognition at ICSI. He's also building a virtual machine that provides access through one simple interface to all the tools commonly used by ICSI speech researchers, including general tools, like the Hidden Markov Model Toolkit, and those developed at ICSI, like Relative Spectral Processing. He will be working remotely from Spokane, Washington.

"I'm really happy to be back," he said.

speech group highlights

Researchers will work with **Microsoft Corp.** to advance the state of the art in human-computer interaction that relies on speech and other modalities. Elizabeth Shriberg and Andreas Stolcke, principal scientists with Microsoft's Conversational Systems Laboratory and ICSI external fellows, will lead the effort. In one of the first projects, researchers will use information conveyed by the melody and rhythm of speech, known as prosody, to improve automatic speech understanding.

Researchers from **Nanyang Technical University** (NTU) in Singapore will visit the Speech Group over the next two years. Beginning September 1, NTU will send researchers from its Temasek Laboratories to visit ICSI in order to collaborate on speech research techniques. In the first year of the agreement, NTU researchers will work on multilingual speech recognition and keyword search, and in the second, on speech recognition.

new projects

spoken wordsearch with rapid development and frugal invariant subword hierarchies

(SWORDFISH)
PI: Nelson Morgan
Funded by: IARPA

Researchers develop ways to find spoken phrases in audio from multiple languages, using a fraction of the training data and time usually required for such a task.

towards modeling human speech confusions in noise

PI: Nelson Morgan
Funded by: NSF

Researchers study how background noise and speaking rate affect the ability of humans to recognize speech by evaluating components of a model of human speech perception.

outing unfortunate characteristics of HMMs

(OUCH)
PIs: Nelson Morgan. Steven Wegmann, Jordan Cohen
Funded by: IARPA

Researchers explore automatic speech recognition to understand the limitations and challenges of current technologies.

news briefs

DAAD-funded postdoctoral fellow **Nils Peters** won a best paper award at the Sound and Music Computing Conference for his work on the Spatial Sound Description Interchanges Format (SpatDIF). Peters's research in signal processing and spatial acoustics aims to analyze and semantically describe fields of sound. He works with an array of 150 microphones built by UC Berkeley's Center for New Music and Audio Technologies and Meyer Sound. The microphone array senses the sounding objects in a room environment, and the captured data are processed by algorithms to identify the room environment and to classify and detect the location of the sounding objects. SpatDIF organizes these sound field descriptions in a structured way. Peters was presented with the award on July 14.

Faculty associate **Christos Papadimitriou** and his collaborator are one of three teams sharing the 2012 Gödel Prize. Papadimitriou and Elias Koutsoupias are recognized for their paper "Worst-Case Equilibria," which examines the effect of selfish networking behavior. The prize is awarded annually for outstanding papers in theoretical computer science by ACM's Special Interest Group on Algorithms and Computation Theory and the European Association for Theoretical Computer Science. It was presented on July 12 at the International Colloquium on Automata, Languages and Programming. Papadimitriou is the C. Lester Hogan Professor of Electrical Engineering and Computer Science at UC Berkeley.


Professor Charles Fillmore

Professor **Charles Fillmore**, director of the FrameNet Project, received the Association for Computational Linguistics's Lifetime Achievement Award on July 11. The award is given for widely recognized, sustained, and enduring contributions to the field of computational linguistics over a long period. Fillmore has been and continues to be a major contributor to contemporary linguistics. The FrameNet Project is based on the theory of frame semantics, which Fillmore developed in the 1970s. Fillmore also developed case grammar theory in the 1970s, and his work in collaboration with **Paul Kay** and **George Lakoff** was generalized into the theory of construction grammar.

Michael Luby, a long-time collaborator and the former leader of the Algorithms Group, received the IEEE Richard W. Hamming Medal on June 30. Luby shares the award with alum Amin Shokrollahi. They are recognized for developing efficient and flexible data coding methods that have enabled the success of information distribution applications such as video streaming and delivery of data to mobile devices. Luby is the vice president of technology at QUALCOMM's office in Berkeley, and Shokrollahi is a professor of math and computer science at EPFL and the former chief scientist of Digital Fountain in Lausanne, Switzerland. He is also the CEO of Kandou Bus, a company he co-founded in 2011.

Bro 2.1 was publicly released August 29, shortly after the 2012 Bro Exchange, a meeting that brought together Bro users to exchange thoughts on and experiences with deploying the system. Bro is an open-source network security monitor developed by a team of researchers and engineers at ICSI and the National Center for Supercomputing Applications at the University of Illinois. Bro monitors network traffic at major universities, large research labs, supercomputing centers, and open–science communities around the country. It has been the cornerstone of network security at the Lawrence Berkeley National Laboratory since the 1990s. Recent installations include an OpenFlow-based Bro cluster setup at Indiana University. The University of Utah's Center for High Performance Computing is also currently investigating how to integrate the monitor into its new security architecture. The new 2.1 release introduces extensive support for IPv6, tunnel decapsulation, a new input framework for integrating external information in real-time into the processing, and two new experimental log output formats.

Director **Roberto Pieraccini** has been named a 2012 speech luminary by *Speech Technology* magazine. The speech luminary awards are given each year to honor those whose creativity and drive have significantly influenced the speech technology industry. Pieraccini is one of five people honored this year. He is cited, among other things, for his contributions to spoken language understanding and for his book, *The Voice in the Machine*, "hailed as a wake-up call to the speech industry, which Pieraccini concludes has yet to master human-computer interactions."

scott shenker
to serve as chief scientist

ICSI has appointed Scott Shenker as its first chief scientist. Shenker, who has led the Networking Group since 1998, will help set its research agenda and represent the interests of its researchers.

"We established the role of chief scientist at ICSI to ensure the Institute maintains its research excellence and expands its worldwide scientific relevance in today's quickly evolving technological landscape," said Director Roberto Pieraccini. "And Scott Shenker, a true visionary, a role model in the computer science community, and one of the pillars of ICSI, is ideally qualified to take on this role."

Shenker is well known for his contributions to Internet architecture and other topics in network design. He is also a leader, with Martin Casado and Nick McKeown, in the movement toward software-defined networking. He serves on the board of the Open Networking Foundation. Earlier this year, Shenker was elected a member of the National Academy of Engineering; he is also a fellow of the ACM and the IEEE and has received the ACM SIGCOMM and IEEE Internet Awards.

"Scientists are the lifeblood of any research organization. As chief scientist, I have the honor of representing their interests at ICSI as we plan for the future," Shenker said. "I'm very excited about this opportunity to help ICSI expand its research agenda to address new scientific challenges."

networking group to study human element of cybercrime

ICSI, the University of California, San Diego, and George Mason University have received a $10 million, five-year grant from the National Science Foundation to investigate the roles played by economics and social interactions in Internet security.

"During our earlier work on analyzing the factors that go into making spam a profitable form of cybercrime, we were deeply struck by the significance of the human side of the equation," said Vern Paxson of the Networking Group, one of the leaders of the project and a UC Berkeley professor. "Non-technical considerations span business concerns, issues of trust-amongst-thieves, and the rise of social media as both a new domain that cybercrime is expanding into, and a way to track interactions amongst the criminals themselves."

Security research has tended to focus on the technologies that enable and defend against attacks. The new project, led by Paxson and Stefan Savage of UC San Diego, emphasizes the profits that motivate the majority of Internet attacks, the elaborate marketplaces that support them, and the relationships among cybercriminals, who rely upon each other for services and expertise.

Paxson and Savage have a long history of working together on Internet security. In 2004, they established the NSF-funded Center for Internet Epidemiology and Defenses (CCIED), which they have led since then. More recently, they've led a large-scale effort funded by the Office of Naval Research on infiltrating the "botnets," groups of malware-infected computers under the command of a single person, that attackers often use for their attacks. Recent results from CCIED include findings that just three banks authorized 95 percent of credit card sales of goods advertised through spam. This kind of research aims to help the fight against spam and malware by exposing weak points in the spam economic chain.

The new project, a natural outgrowth of this, focuses on the human element of cybercrime, including how social media such as Facebook and Twitter provide new opportunities for attacks and manipulation. By better understanding the roles that economics and social interactions play in cybercrime, the researchers say, defenders can identify the most effective opportunities for interventions and defenses.

visiting scholars
june 2012 – september 2012

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

AI
Oliver Culo (Germany)
Gerard de Melo (Germany)
Sergio Guadarrama
Emanuel Kitzelmann (Germany)
Hiroaki Sato
Malte Schilling (Germany)

Algorithms
Itamar Eskin
Farhad Hormozdiari
Ron Shamir

Architecture
Miquel Moretó Planas

Vision
Tim Althoff
Tobias Baumgartner
Stefanie Jegelka
Matthias Kirchner (Germany)
Lorenzo Riano

Networking
Bernhard Amann (Germany)
Soumya Basu
Jason Croft
Lorenzo DeCarli
Mohan Dhawan
Haixin Duan
Oana Goga
Shuang Hao
Sheharbano Khattak
Eemil Lagerspetz (Finland)
Amin Tootoonchian
Sam Whitlock
Andreas Wundsam (Germany)

Speech
Seppo Enarvi (Finland)
Arlo Faria (Brazil)
Frantisek Grezl
Mikko Kurimo (Finland)
David Suendermann

Campus Affiliation / Other
Nils Peters (Germany)

Seppo Enarvi

Stefanie Jegelka

publications

Y. Baran, B. Pasaniuc, S. Sankararaman, D. G. Torgerson, C. Gignoux, C. Eng, W. Rodriguez-Cintron, R. Chapela, J. G. Ford, P. C. Avila, J. Rodriguez-Santana, E. G. Burchard, and E. Halperin. *Fast and Accurate Inference of Local Ancestry in Latino Populations. Bioinformatics, Vol. 28, Issue 10, pp. 1359-1367, May 2012.*

M. Casado, T. Koponen, S. Shenker, and A. Tootoonchian. *Fabric: A Retrospective on Evolving SDN. Proceedings of the ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN), Helsinki, Finland, August 2012.*

J. Choi, G. Friedland, V. Ekambaram, and K. Ramchandran. *Multimodal Location Estimation of Consumer Media – Dealing with Sparse Training Data. Proceedings of the IEEE International Conference on Multimedia and Expo, Melbourne, Australia, pp. 43-48, July 2012.*

O. Culo and G. de Melo. *Source-Path-Goal: Investigating the Cross-Linguistic Potential of Frame-Semantic Text Analysis. it - Information Technology, Vol. 54, Issue 3, pp. 147-152, May 2012.*

G. de Melo, C. F. Baker, N. Ide, R. J. Passonneau, and C. Fellbaum. *Empirical Comparisons of MASC Word Sense Annotations. Proceedings of the 8th Conference on International Language Resources and Evaluation (LREC 2012), Istanbul, Turkey, May 2012.*

G. de Melo and G. Weikum. *UWN: A Large Multilingual Lexical Knowledge Base. Proceedings of the 50th Annual Meeting of the Association of Computational Linguistics (ACL 2012), Jeju, Korea, pp. 151-156, July 2012.*

A. Efros and E. Halperin. Haplotype Reconstruction Using Perfect Phylogeny and Sequence Data. Proceedings of the Second Annual Satellite Workshop on Massively Parallel Sequencing at the 16th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2012), Barcelona, Spain. Published in BMC Bioinformatics, Vol. 13, Suppl. 6:S2, April 2012.

G. Friedland. Cybercasing the Joint: Language Technologies, Multimedia Retrieval, and Online Privacy. Presented at the Language Technologies Institute Colloquium, Carnegie Mellon University, Pittsburgh, Pennsylvania, April 13, 2012.

G. Friedland, B. Martinez Elizalde, H. Lei, and A. Divakaran. There is No Data Like Less Data: Percepts for Video Concept Detection on Consumer-Produced Media. ICSI Technical Report TR-12-006, April 2012.

O. Goga, H. Lei, S. H. K. Parthasarathi, G. Friedland, R. Sommer, and R. Teixeira. On Exploiting Innocuous User Activity for Correlating Accounts Across Social Network Sites. ICSI Technical Report TR-12-008, May 2012.

N. Ide, C. Baker, C. Fellbaum, and R. Passonneau. The MASC/MultiMASC Community Collaboration Project: Why You Should Be Involved and How. Proceedings of the Workshop on Collaborative Resource Development and Delivery at the 8th Conference on International Language Resources and Evaluation (LREC 2012), Istanbul, Turkey, May 2012.

Y. Jia, C. Huang, and T. Darrell. Beyond Spatial Pyramids: Receptive Field Learning for Pooled Image Features. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), Providence, Rhode Island, pp. 3370-3377, June 2012.

R. M. Karp. Theory of Computation as an Enabling Tool for the Sciences. Presented at the Turing Centennial Celebration, Princeton, New Jersey. Also presented at the 9th Annual Conference on Theory and Applications of Models of Computation (TAMC 2012), Turing Centenary Meeting, Beijing, China, May 2012.

H. Lei, J. Choi, and G. Friedland. Multimodal City-Identification on Flickr Videos Using Acoustic and Textual Features. ICSI Technical Report TR-12-007, May 2012.

D. Levin, A. Wundsam, B. Heller, N. Handigol, and A. Feldmann. Logically Centralized? State Distribution Tradeoffs in Software Defined Networks. Proceedings of the Workshop on Hot Topics in Software Defined Networks (HotSDN '12) at ACM SIGCOMM 2012, Helsinki, Finland, August 2012.

D. McCoy, A. Pitsillidis, G. Jordan, N. Weaver, C. Kreibich, B. Krebs, G. M. Voelker, S. Savage, and K. Levchenko. PharmaLeaks: Understanding the Business of Online Pharmaceutical Affiliate Programs. Proceedings of the 21st USENIX Security Symposium, Bellevue, Washington, August 2012.

A. J. Oliner, A. P. Iyer, E. Lagerspetz, I. Stoica, and S. Tarkoma. Carat: Collaborative Energy Bug Detection. Poster and demo at the proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI '12), San Jose, California, April 2012.

R. Passonneau, C. Baker, C. Fellbaum, and N. Ide. The MASC Word Sense Corpus. Proceedings of the 8th Conference on International Language Resources and Evaluation (LREC 2012), Istanbul, Turkey, pp. 3025-3030, May 2012.

N. Peters, T. Lossius, and T. Place. An Automated Testing Suite for Computer Music Environments. Proceedings of the 9th Sound and Music Computing Conference (SMC 2012), Copenhagen, Denmark, July 2012

N. Peters and S. McAdams. A Perceptual Analysis of Off-Center Sound Degradation in Surround-Sound Reproduction Based on Geometrical Properties. Journal of the Acoustical Society of America, Vol. 131, Issue 4, pp. 3256-3256, April 2012. Also in the proceedings of Acoustics 2012, Hong Kong, May 2012.

N. Peters, J. Schacher, and T. Lossius. SpatDIF: Principles, Specification, and Examples. Proceedings of the 9th Sound and Music Computing Conference (SMC 2012), Copenhagen, Denmark, July 2012.

M. R. L. Petruck and G. de Melo. Precedes: A Semantic Relation in FrameNet. Proceedings of the Workshop on Language Resources for Public Security Applications at the 8th Conference on International Language Resources and Evaluation (LREC 2012), Istanbul, Turkey, pp.45-49, May 2012.

B. Raghavan, D. Irwin, J. Albrecht, J. Ma, and A. Streed. An Intermittent Energy Internet Architecture. Proceedings of the Third International Conference on Future Energy Systems (e-Energy 2012), Madrid, Spain, May 2012.

C. Rossow, C. J. Dietrich, C. Kreibich, C. Grier, V. Paxson, N. Pohlmann, H. Bos, and M. van Steen. Prudent Practices for Designing Malware Experiments: Status Quo and Outlook. Proceedings of the 33rd IEEE Symposium on Security and Privacy (S&P 2012), San Francisco, California, pp. 65-79, May 2012.

R. Rozov, E. Halperin, and R. Shamir. MGMR: Leveraging RNA-Seq Population Data to Optimize Expression Estimation. Proceedings of the Second Annual Satellite Workshop on Massively Parallel Sequencing at the 16th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2012), Barcelona, Spain. Published in BMC Bioinformatics, Vol. 13, Suppl. 6:S2, April 2012.

M. Sargent, B. Stack, T. Dooner, and M. Allman. A First Look at 1 Gbps Fiber-To-The-Home Traffic. ICSI Technical Report TR-12-009, August 2012.

R. Sharan and R. M. Karp. Reconstructing Boolean Models of Signaling. Proceedings of the 16th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2012), Barcelona, Spain, pp. 261-271, April 2012.

S. Shenker. Software-Defined Networking: History, Hype, and Hope. Proceedings of the Conference on Large-Scale Distributed Systems and Middleware (LADIS 2012), Madeira, Portugal, July 2012. Also in the proceedings of the Brazilian Symposium on Computer Networks and Distributed Systems (SBRC 2012), Ouro Preto, Brazil, April 2012.

C. Shue, A. Kalafut, M. Allman, and C. Taylor. On Building Inexpensive Network Capabilities. ACM SIGCOMM Computer Communications Review, Vol. 42, No. 2, pp. 72-79, April 2012.

K. Thomas, C. Grier, and V. Paxson. Adapting Social Spam Infrastructure for Political Censorship. Proceedings of the Fifth USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET '12), San Jose, California, April 2012.

A. Tootoonchian, S. Gorbunov, Y. Ganjali, M. Casado, and R. Sherwood. On Controller Performance in Software-Defined Networks. Proceedings of the USENIX NSDI Workshop on Hot Topics in Management of Internet, Cloud, and Enterprise Networks and Services (Hot-ICE), San Jose, California, April 2012.

Z. Wang, F. Hormozdiari, W.-Y. Yang, E. Halperin, and E. Eskin. CNVeM: Copy Number Variation Detection Using Uncertainty of Read Mapping. Proceedings of the 16th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2012), Barcelona, Spain, pp. 326-340, April 2012.

W.-Y. Yang, J. Novembre, E. Eskin, and E. Halperin. A Model-Based Approach for Analysis of Spatial Structure in Genetic Data. Nature Genetics, Vol. 44, pp. 725-731, May 2012.

O. Zagordi, A. Topfer, S. Prabhakaran, V. Roth, E. Halperin, and N. Beerenwinkel. Probabilistic Inference of Viral Quasispecies Subject to Recombination. Proceedings of the 16th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2012), Barcelona, Spain, pp.342-354, April 2012.

N. Zhang, R. Farrell, and T. Darrell. Pose Pooling Kernels for Sub-Category Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), Providence, Rhode Island, pp. 3665-3672, June 2012.

W. Zhang, S. X. Yu, and S.-H. Teng. Power SVM: Generalization with Exemplar Classification Uncertainty. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), Providence, Rhode Island, pp. 2144-2151, June 2012.

INTERNATIONAL COMPUTER SCIENCE INSTITUTE

THE INTERNATIONAL COMPUTER SCIENCE INSTITUTE (ICSI) is one of the few independent, non-profit US labs solely conducting open, non-proprietary, pre-competitive research in computer science. ICSI's mission is to further research in computer science through international collaboration, and further international collaboration through research in computer science. Affiliated with the University of California campus in Berkeley, 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 vision; and computer architecture.

Editor-in-Chief
Maria Quintana
Senior Editor
Roberto Pieraccini
Design
envision ID
Copywriter
Chris Switzer
Layout
Jacob Wolkenhauer
Photography
Jaci Considine
Leah Hitchcock

INTERNATIONAL
COMPUTER
SCIENCE
INSTITUTE

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

facebook.com/ICSIatBerkeley

twitter.com/ICSIatBerkeley