COMPUTING edge

> Mobile Video Delivery: Challenges and Opportunities

> Multimedia Search: From Relevance to Usefulness

> Biometric Authentication on Mobile Devices

JULY 2015

www.computer.org

IEEE

IEEE computer society

STC 2015

27th Annual IEEE Software Technology Conference (STC 2015)

STC

IEEE SOFTWARE TECHNOLOGY CONFERENCE

Call for Presentations

October 12-15, 2015

Hilton Long Beach, Long Beach, California, USA

http://ieee-stc.org

Theme: Meeting Real World Opportunities and Challenges through Software and Systems Technology

The IEEE Software Technology Conference (STC) is a leading software conference that targets software practitioners from industry, government, and academia. The conference provides a forum for leading software practitioners to share opportunities, challenges, and successes. The rich exchange of ideas, technologies, data, and best practices will help practitioners lead software for business innovation, national defense, critical infrastructure, information technology, global markets, etc.

We invite you to connect with colleagues and share your knowledge, expertise, and insights in software and systems technology for meeting real world opportunities and challenges. Presenters or tutorial leaders should please submit their abstracts using this website **http://ieee-stc.org**. Presenters have the option to submit full papers instead of abstracts, and all accepted full papers will be published in the IEEE Xplore Digital Library.

Focus areas include:

- *Systems engineering for software-intensive systems*
- *Software architectures*
- *Agile development*
- *Cybersecurity*
- *Analytics*
- *Mobile applications*
- *Cloud computing*
- *Open source*
- *Supply chain risk management*
- *Critical infrastructure challenges*
- *Affordability*
- *Human capital and knowledge transition*

Important Dates

July 15, 2015	Submit abstracts and optional full papers
July 31, 2015	Send acceptance notifications
August 21, 2015	Submit final abstracts and presenter biographies
September 25, 2015	Submit final presentations and optional full papers

IEEE

IEEE computer society

COMPUTING edge

IEEE COMPUTER SOCIETY http://computer.org • +1 714 821 8380

Printed with inks containing soy and/or vegetable oils

IEEE

IEEE computer society

BPA WORLDWIDE

Circulation: ComputingEdge is published monthly by the IEEE Computer Society. IEEE Headquarters, Three Park Avenue, 17th Floor, New York, NY 10016-5997; IEEE Computer Society Publications Office, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720; voice +1 714 821 8380; fax +1 714 821 4010; IEEE Computer Society Headquarters, 2001 L Street NW, Suite 700, Washington, DC 20036.

Postmaster: Send undelivered copies and address changes to IEEE Membership Processing Dept., 445 Hoes Lane, Piscataway, NJ 08855. Application to Mail at Periodicals Postage Prices is pending at New York, New York, and at additional mailing offices. Canadian GST #125634188. Canada Post Corporation (Canadian distribution) publications mail agreement number 40013885. Return undeliverable Canadian addresses to PO Box 122, Niagara Falls, ON L2E 6S8 Canada. Printed in USA.

JULY 2015 • VOLUME 1, NUMBER 7

COMPUTING edge

36

Parallel Coordinates for Multidimensional Data Visualization

Departments

SPOTLIGHT ON TRANSACTIONS

Distributed Optimal Channel Access in Cognitive Radio Networks

Prasant Mohapatra, University of California, Davis

This installment highlighting the work published in IEEE Computer Society journals comes from IEEE Transactions on Mobile Computing.

Cognitive radio networks (CRNs) let unlicensed or secondary users (SUs) access licensed spectra, improving wireless resource utilization and addressing the problem of spectrum shortage. SUs are allowed spectrum access only when they don't cause unacceptable levels of interference to licensed or primary users (PUs). Although throughput-optimal algorithms based on the well-known maximal-weight scheduling algorithm exist for CRNs, they require central processing of network-wide SU information. Developing a distributed implementation that can fully leverage the spectrum opportunities for SUs has so far remained elusive.

Because distributed algorithms eliminate the need for network-wide state information collection and avoid costly algorithm executions, researchers Shuang Li, Eylem Ekici, and Ness Shroff developed a fully distributed channel-access solution that uses only local queue-length information.

In their article "Throughput-Optimal Queue Length–Based CSMA/CA Algorithm for Cognitive Radio Networks" (*IEEE Trans. Mobile Computing*, vol. 14, no. 5, 2015, pp. 1098—1108), the authors identify a new class of *queue length–based carrier sense multiple access* (CSMA) algorithms as a promising approach to solving the distributed and optimal channel-access problems in CRNs. This new class of random-access algorithms combines channel-sensing results with locally available queue-length information to determine channel-access probabilities, achieving full capacity in ad hoc wireless networks in a distributed manner. However, the throughput optimality is realized for nonfading and always-available channel cases, a condition not met in CRNs due to PU activity.

Li, Ekici, and Shroff overcome this problem with a novel algorithm based on a new system-state representation that tracks the channel-availability information. This new definition of the system state, along with the differentiated treatment of SUs inside and outside the PU interference range, leads to a distributed algorithm design that achieves provable throughput optimality for SUs without disrupting PU operation. Significant capacity gains realized by the algorithm vis-à-vis other queue length–based CSMA algorithms highlight the efficacy of the authors' design approach.

The researchers' results show that it's possible to garner the full potential of CRNs without relying on complex centralized algorithms. This work is the first step toward highly efficient, decentralized deployment of CRNs and will lead to the development of other distributed optimal algorithms for networks with time-varying channel conditions.

PRASANT MOHAPATRA is a professor in the Department of Computer Science at the University of California, Davis, and the editor in chief of *IEEE Transactions on Mobile Computing*. Contact him at pmohapatra@ucdavis.edu.

 2376-113X/15/$31.00 © 2015 IEEE

Magazine Roundup

The IEEE Computer Society's lineup of 13 peer-reviewed technical magazines covers cutting-edge topics ranging from software design and computer graphics to Internet computing and security, from scientific applications and machine intelligence to cloud migration and microchip manufacturing. Here are highlights from recent issues.

Computer

As companies develop wearable computers, people are increasingly utilizing the devices to provide or augment human capabilities. In the process, they are using them in new ways. This and other matters are discussed in *Computer*'s June 2015 special issue on **wearable computing**.

IEEE Software

During its 10-year history, the annual Software Engineering Institute Architecture Technology User Network (SATURN) conference has become a barometer for the ever-evolving world of software architecture. "Lightweight and Flexible: Emerging Trends in Software Architecture from the SATURN Conferences"—from *IEEE Software*'s May/June 2015 issue—summarizes **software architecture trends** that emerged during SATURN 2014 and looks to the future.

IEEE Internet Computing

In the past decade, social-networking services have flooded the Web. The public nature of such services makes **identity deception** easy. Fighting this problem requires a coordinated approach by users and developers. "Detecting and Preventing Online Identity Deception in Social Networking Services," from *IEEE Internet Computing*'s May/June 2015 issue, identifies and evaluates the most prevalent approaches.

Computing in Science & Engineering

Public-funding bodies in the UK and US have released documents outlining a similar vision for **funding e-infrastructure development** to support scientific collaboration. These documents have strengths and weaknesses, which are discussed in "The Future of E-Infrastructures," in *CiSE*'s May/June 2015 issue.

IEEE Security & Privacy

IEEE S&P's May/June 2015 special issue, titled "**Security on Tap**," features articles on improving security with diversity, better cryptographic protocols, and a new pervasive-computing identity-management model. It also highlights India's cybersecurity challenges and examines different types of attacks.

IEEE Cloud Computing

Engineering the Internet of Things (IoT) and cloud services to provide a coherent software layer for continuous deployment, provision, and execution of applications for various domains is complex. The authors of "Principles for Engineering IoT Cloud Systems," from *IEEE Cloud Computing*'s March/April 2015 issue, consider whether and how IoT cloud systems could provide a uniform layer to enable continuous execution of complex applications consisting of diverse software components.

IEEE Computer Graphics and Applications

Robots' increasing popularity makes interacting and communicating with them of critical importance. *IEEE CG&A*'s May/June 2015 special issue—"Natural User Interfaces for Adjustable Autonomy in Robot Control"—presents some of the latest research results in the exploration of **natural user interfaces for robotic control and teleoperation**.

IEEE Intelligent Systems

IEEE Intelligent Systems' March/April 2015 issue features articles on macro-level predictive analytics, which has benefited from the ability to use big data for real-time analysis. The magazine's May/June 2015 special issue—titled "Predictive Analytics: Predictive Modeling at the Micro Level"—addresses **micro-level predictive analytics**, which uses data to make inferences about unknown outcomes relating to individual firms, people, or situations.

IEEE MultiMedia

The volume and nature of **YouTube-generated traffic** has had a significant impact on fixed and mobile networks. Studying YouTube traffic characteristics could enable network engineers to come up with ways to regulate network traffic and support the development of sustainable video-delivery services. In "A Survey of Current YouTube Video Characteristics," in *IEEE MultiMedia*'s April–June 2015 issue, the authors discuss their study.

IEEE Annals of the History of Computing

Few areas of computing capture more headlines than cybersecurity. However, historians have barely scratched the surface of **computer security's history**, largely because of the lack of relevant source materials. Recently, this has begun changing, as researchers have uncovered archival information and developed oral histories. Some of this work is addressed in *IEEE Annals*' April–June 2015 special issue on cybersecurity.

IEEE Pervasive Computing

"**Approximate Computing**: Making Mobile Systems More Efficient," in *IEEE Pervasive Computing*'s April–June 2015 issue, looks at technology that makes systems faster and more efficient by being less than perfect for tasks that don't require great precision. The article explores work in this area by University of Washington researchers.

IT Professional

Two increasingly high-profile technology concepts are sustainability and the Internet of Things (IoT), also known as the **Internet of Anything**. These concepts are evolving and beginning to converge in some cases, creating interesting possibilities. *IT Pro*'s May/June 2015 special issue covers the Internet of Anything and sustainability.

IEEE Micro

IEEE Micro's March/April 2015 special issue highlights some of the processors presented at 2014's Hot Chips 26 conference. One of those chips is discussed in the article "Sparc64 XIfx: Fujitsu's **Next-Generation Processor for High-Performance Computing**." The chip is designed for massively parallel supercomputer systems.

Computing Now

The Computing Now website (http://computingnow.computer.org) features **up-to-the-minute computing news** and blogs, along with articles ranging from peer-reviewed research to opinion pieces by industry leaders.

EDITOR'S NOTE

The Vibrant World of Graphics

This issue of *ComputingEdge* explores the exciting, high-profile world of computer graphics.

Graphics technology for creating still images, animation, and video has grown and matured over the years. New approaches and techniques have improved both the process and the results, yielding more appealing images, games, and movie special effects. The *IEEE Computer Graphics and Applications (CG&A)* article "Electrifying Digital Abstract Art" features an artist who uses cutting-edge technology to create digital works.

However, graphics has become more than just a way to create multimedia for entertainment. For example, applications are using graphics to generate easy-to-grasp visual representations of patterns in large datasets. This has become critical to the growing field of analytics, particularly big-data analytics. Various aspects of this topic are addressed in *Computing in Science and Engineering*'s "Parallel Coordinates for Multidimensional Data Visualization" and *CG&A*'s "Preparing Undergraduates for Visual Analytics."

Graphics can also enhance the educational experience. *IEEE MultiMedia*'s "Teaching Privacy: Multimedia Making a Difference" examines a research program that has developed apps for teaching young people to make informed choices about online privacy.

Meanwhile, graphics' popularity has created numerous peripheral issues that users encounter, such as how to deal with with large image collections. *MultiMedia*'s "Photos to Remember, Photos to Forget" highlights a European project focusing on technology that can help people decide which content is important enough to preserve and which isn't.

Another issue is that people want to watch video on their smartphones and tablets, but providing data-intensive high-definition video over resource-constrained mobile networks can be difficult, as noted in *IEEE Internet Computing*'s "Mobile Video Delivery: Challenges and Opportunities."

Many users want the ability to search multimedia collections in the same way and with the same precision that they conduct text-based Internet searches. Researchers are looking for ways to provide this capability, as covered in *MultiMedia*'s "Multimedia Search: From Relevance to Usefulness."

ComputingEdge articles on topics other than graphics include:

- *IEEE Software*'s "Lightweight and Flexible," which looks at software-architecture trends discussed during 2014's Software Engineering Institute Architecture Technology User Network (SATURN) conference;
- *IEEE Intelligent Systems*' "Simple is Beautiful: Toward Light Prediction Markets," which examines intelligent predictive technology;
- *IEEE Security and Privacy*'s "Biometric Authentication on Mobile Devices"; and
- *IT Professional*'s "Internet of Things: Making the Hype a Reality."

Electrifying Digital Abstract Art

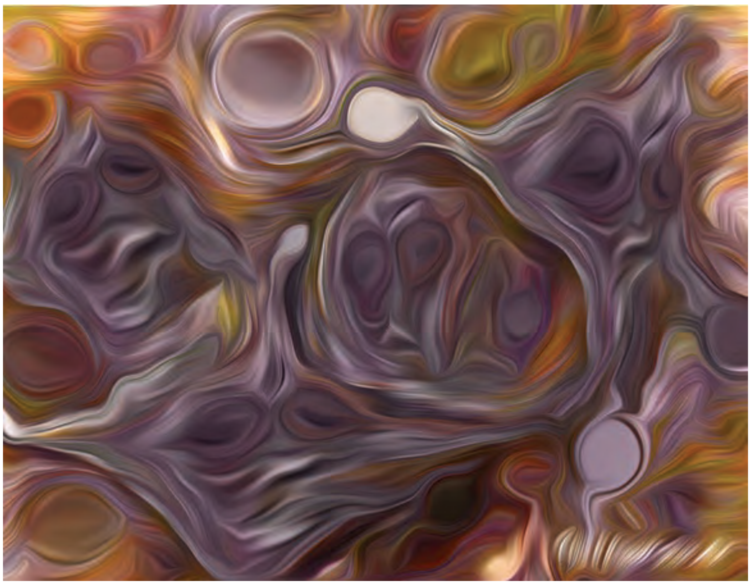
Gary Singh

Merging the analog with the digital, fusing fine art with a design sensibility, Kevin O'Connell never completely leaves any of those realms. He's both a fine artist who shoots representational imagery from the natural world and an electrifying digital abstractionist.

Even better, "electrifying" is not just a metaphor here, as O'Connell really is an electrician. He entered a four-year apprenticeship program straight out of high school—not normally a path applicable to a fine arts or design career. Nevertheless, being an electrician allowed him to travel all over the world, which subsequently allowed him to photograph numerous obscure locations off the beaten path. One pursuit led to the other, so to speak.

"All the men in my family including grandfathers, father, uncles, and cousins were electricians, so I was expected to do the same," O'Connell recalled. "This job was never appealing to me, so I also freelanced as a photographer shooting weddings and portraits as well. In 2005, I needed to work as an artist, photographer full time, so I decided to use my skill as an electrician to work in remote places around the world that other photographers could not go to unless they had a skilled trade license."

Figure 1. ***Nomads*. The ghost nomads in this piece gradually emerged as O'Connell improvised using layers in Photoshop.**



As a result, O'Connell lived all over the globe for five years, finding himself in places like Antarctica or islands in the central Pacific Ocean. High school electives in traditional photography, previously thought to only be a hobby, now came to the forefront and O'Connell found his calling. A few years ago, he went back to college to pursue a digital design degree at the Illinois Institute of Art, which only further enhanced his digital aesthetic.

"Digital imagery made it fun and exciting to design abstract-style digital paintings with a twist," O'Connell says. "I have always been very fond of colors and shadow, so working with digital imagery I can change the shadow structure of individual pieces and parts along with different depths of color."

Ghosts and Kitchens

O'Connell pulled an all-nighter when working on *Buttons and Lace* (the cover image). The painting flowed every step of the way, through various Photoshop incarnations.

"I started with an idea of taking round circles and placing them in a row, and then wanted to give them contour and texture to set them apart," O'Connell explains. "From there, I would sit back and study the piece and add to it with different contrasts and shapes."

O'Connell says he prefers bold colors, and since color plays a huge role in his designs, every portion of the image needed to be a certain hue. Then a few textures gave the image some much-needed depth.

"When starting a new digital art, I usually don't have any particular color in mind. I start with a couple different colors, and as I go forward, I change them periodically to enhance different sections or parts or to give it more or less contrast."

With *Nomads* (see Figure 1), the process was more improvisatory than his usual approach. O'Connell tweaked away in the software until various interlocking ghost-like faces began to emerge, triggering memories of a film in which nomadic ghosts floated around their environment, wreaking havoc among the inhabitants of a particular landscape.

Published by the IEEE Computer Society 2376-113X/15/$31.00 © 2015 IEEE

"I just wanted to ad lib on this piece and wait for my inspiration to come," he says. "After many hours with layers in Photoshop, I stepped back to look over my work. Deep inside was an outline of a face that stood out to me. I remembered a movie I saw about these deadly ghost nomads and decided to design the rest of this work around these nomads."

Another Dimension is a more representational image, but abstract ethereal elements come into play, nevertheless (see Figure 2). Multidimensional streetscapes merge with mysterious emotive hues of yellow and orange.

"Sometimes representational works can be very intriguing, with many different dimensions within a single piece of art," O'Connell says. "Making these all blend together is quite hard, but when you get it right, they make for very interesting works."

But O'Connell's work is not limited to mystery and intrigue. His design degree especially came in handy when starting his kitchen art operation. Designed to be printed on metal, *Geometrics* came to O'Connell's mind when his parents asked him to supply a traditional painting for their kitchen (see Figure 3). After doing so, he eventually changed his mind and came up with another idea. The kitchen featured cool blue and white components with stainless steel accents, and O'Connell envisioned a piece of metal art that complimented the aspects of the kitchen.

"Kitchens are the most expensive part of a home, and usually the most interesting architecturally," he says. "They rarely have a piece of art solely designed to compliment them either, except for images of fruit, coffee, and other food related ideas. I wanted to design custom stand-alone pieces of art that would really pop, yet be cohesive."

So he took down the traditional painting he originally gave them and designed *Geometrics* in its place. After printing it on metal, he put it up in his parents' kitchen. He now specializes in metal art that matches kitchen décor.

The Fine Art of Design

Now that he's a fine artist and commercial designer, O'Connell gets to navigate both of those landscapes. But he remains skeptical about the two ever completely merging. Traditional brick-and-mortar institutions rarely accept purely digital imagery, and they often view digital prints as reproductions.

"I believe that a digital painting can take just as long or longer that a traditional painting," O'Connell says. "You need a lot of the same skill set for both. However, I do also feel that an original piece of art now becomes interesting. There is no original in a digital piece of art, and until there is a way that can happen, I do agree somewhat with the reasoning of traditional brick-and-mortar institutions."

Figure 2. ***Another Dimension*. A more representational image, this piece features multidimensional streetscapes.**

Nevertheless, O'Connell says a more realistic way to fuse the variety of backgrounds informing his practice would be to design a digital darkroom with electrical automation of lighting movement.

"I would like to incorporate designs from the digital darkroom with computer design to make some very interesting new style abstracts," he says.

Gary Singh *lives and writes in San Jose, California. Contact him at gsingh@avantguild.com.*

This article originally appeared in IEEE Computer Graphics and Applications, *vol. 35, no. 2, 2015.*

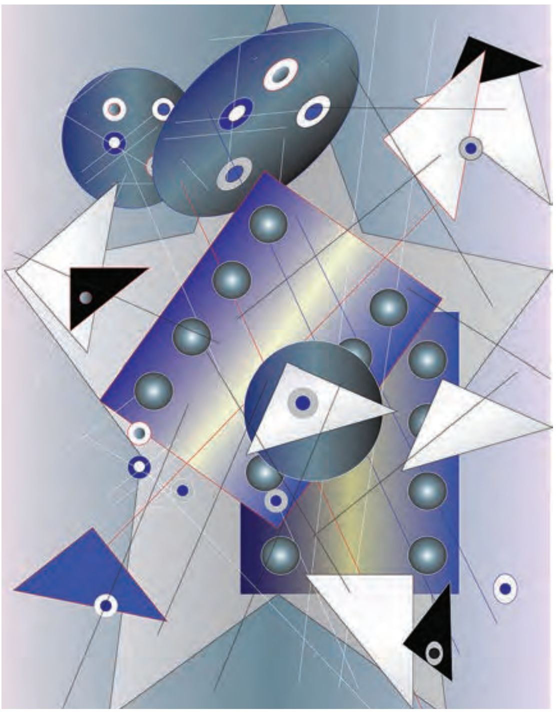


Figure 3. ***Geometrics*. Designed to be printed on metal, this piece came to O'Connell when he was asked to supply a traditional painting for a kitchen.**

Hayley Hung
Technical University of Delft

Teaching Privacy: Multimedia Making a Difference

Julia Bernd, Blanca Gordo, Jaeyoung Choi, Bryan Morgan, Nicholas Henderson, Serge Egelman, Daniel D. Garcia, and Gerald Friedland
International Computer Science Institute and University of California, Berkeley

"When does the information you fill in on a Web form get submitted to the server? Is it submitted when you click okay, or could it already be submitted before that?" A high school computer science teacher asked one of the coauthors these questions. A student had told the teacher a driving school seemed to have her contact information after she visited their website and started to fill in a form, even though she changed her mind and never clicked any buttons to submit that information.

As experts, we know that the answer to the teacher's questions depends on many factors, in particular whether the school used CGI or servlet technology to host the site or whether it used the more recent Ajax or HTML5. But even a relatively tech-savvy high school teacher might find this seemingly simple question hard to answer. Most high school teachers lack the technical background, and even if they did know the answer, how would they explain it to a teenager? Even if they successfully explained it, there is a further issue: Given that the average human would find it rather difficult to tell whether a site used HTML5 versus CGI scripting, what actionable suggestion could be made to help make Web browsing safer?

Much current multimedia research and development centers around applications with great potential to compromise the privacy of Internet users, directly or indirectly. Multimedia scientists and engineers are developing new methods to automatically identify the people depicted in an image or video, or even the person who uploaded it; detect what's happening in an image or video; and determine where it was recorded. As researchers, we often don't think about the privacy implications of such developments down the road. Of course, scientists and technologists want to continue pursuing these fruitful and interesting avenues for enhancing multimedia analysis and retrieval capabilities, but at the same time, we can mitigate the potential negative effects of those capabilities by using our expertise to educate the public about the effects of the new technology on their privacy.

In this article, we describe the Teaching Privacy project at the International Computer Science Institute (ICSI) and the University of California, Berkeley, in which an interdisciplinary team of researchers and educators are developing educational tools to empower K-12 students and college undergraduates in making informed choices about privacy. We describe our interdisciplinary approach to developing and disseminating engaging, interactive educational apps that demonstrate what happens to personal information on the Internet, with a particular focus on multimedia, and our approach to explaining the underlying social and technical principles in accessible terms.

From Research to Education

Teaching Privacy grew out of several strands of work at ICSI and UC Berkeley. These strands came together as researchers working in different areas realized that the explosion of multimedia content being shared on social media was giving rise to a new need for credible information about online privacy—information based on solid technical knowledge rather than panicked speculation.

One major motivation grew out of theoretical research at ICSI on the privacy implications of multimedia technology, including speaker-matching[1,2] and multimedia-retrieval techniques. For example, while working on multimodal location estimation—automatically identifying where an image or video without geotags was recorded according to its visual and acoustic similarity to geotagged media[3,4]—coauthor Gerald Friedland's Multimedia Research Group at ICSI became aware of how few Internet users (at that time) even realized that the images and

Published by the IEEE Computer Society

Guidelines for Socially Responsible, Inclusive Privacy Education

By Blanca Gordo

The ongoing integration of technological innovations into social structures is resulting in substantial changes in privacy—and even our understanding of what constitutes privacy—that affect everyone in the global society of end users. Harmful effects arise both from a lack of comprehensive consumer privacy protections and a lack of scientifically based educational guidelines on how to teach end users about privacy and contextualize how it is changing.

These issues affect both long-term users of Internet technology and new entrants, who are frequently low-income, less-educated, older, and/or immigrant populations. One key difference between experienced and inexperienced users, however, is that new entrants are more vulnerable to potential dangers online because they lack the systematic knowledge that comes with time and continuous direct experience. New entrants are still learning new technical skills, new ways of doing things, and new social norms.

To address this disparity and to provide a framework for privacy decision makers, it is necessary to develop a comprehensive, cross-disciplinary theoretical concept of privacy. Such a grounded conceptualization must account for, among other things, the interconnectivity of social, institutional, economic, political, and technology systems; the dynamic effects of social context and governance; the function of network communication technology as transferring and connecting past, present, and future bits of information; the nature of information as a commodity; the limitations of the legal system with regard to corporations' treatment of personal information, including undisclosed tracking and third-party data use; the trade-offs between security and privacy in government surveillance; sophisticated advancements in analytics and the availability of big data; and the range of individuals' views and online experience.

Such a conceptual framework can help us build more universally applicable educational messages that can help anyone grasp the trade-offs inherent in Internet use, the common possibilities for harm, and the general ramifications of online behavior. We have much to learn about the most strategic, effective ways to explain how technology design, policy, corporate structure, social behaviors, and values lead to privacy-related outcomes. But understanding how privacy works in people's daily lives, and how a range of populations conceptualize those workings and process information, can inform the design of effective educational resources that resonate with everyone's daily experience.

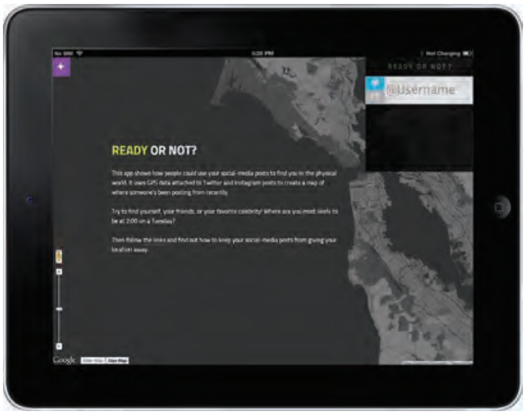
We need to build tools that help users grasp the mechanisms of transmission, collection, storage, and leakage of bits of information that make up someone's personal profile. To accomplish this, digital tools can simulate the vital components of privacy, including technology systems, social behavior, cultural norms, and governance, and illustrate (via metaphors as well as literal descriptions) how technology works within the frames of networked structures and of social-institutional systems.

Taking the experience of new entrants into account when building privacy-related educational tools will also help us to better teach long-term users, who also frequently have misperceptions about privacy and may tend to take it for granted that they can control their personal information. For example, even many experienced users believe that they can remain anonymous if they use an anonymization proxy or that secondary use of personal information must be authorized by the individual. The easier we can make it for the disconnected to grasp the workings of privacy, the easier it will be for us to educate everyone.

videos they uploaded often included GPS metadata in the first place, much less that it was possible to estimate the recording locations of nontagged media. The multimedia group began working with privacy and security researchers at ICSI, including Robin Sommer and Nicholas Weaver, to explore the potential privacy implications of so much multimedia data and metadata being constantly uploaded and shared.[5]

One of the initial sparking moments for the education project was when coauthor Daniel Garcia invited Friedland to talk about this multimedia-privacy research in a professional development session for high school computer science teachers. The teachers found it by far the most engaging topic of the day and asked a multitude of questions—not just because it was technically interesting, but because they were so eager for information about privacy that they could pass on to their students. They were well aware that their students needed more information and guidance about online privacy but felt unqualified to teach about it because they were not themselves sufficiently well-versed in the technical details.

Similarly, coauthor Blanca Gordo's interest arose out of her social science research on developing a theoretical framework for



(a)

(b)

(c)

Figure 1. Ready or Not? interactive app. (a) Welcome screen, (b) a 2013 heatmap and timeline result showing the frequent coordinates of a not entirely random Twitter user, and (c) information about how to turn off location services.

understanding network technology and a pedagogy for teaching technology to people who are newly going online (see the "Guidelines for Privacy Education" sidebar). Those charged with teaching these new entrants about the technology know it is vital to explain the privacy implications of online activities, but they too do not have the necessary expertise and find few ready-made learning materials.[6] Meanwhile, coauthor Serge Egelman's research in human-computer interaction focuses on how people make decisions about their online privacy and how to help them make better ones.

Teaching Privacy also includes the Berkeley Foundation for Opportunities in Information Technology (BFOIT),[7] of which coauthor Nicholas Henderson is a program lead. BFOIT provides historically underrepresented ethnic minority and female middle and high school students with knowledge, resources, practical programming skills, and guidance in their pursuit of higher education and production of technology. BFOIT educators were aware of a deep need for accurate, in-depth privacy education among the program participants—in their current online activities, but especially if they were to go on to become technology designers.

Multimedia Apps and Learning Tools

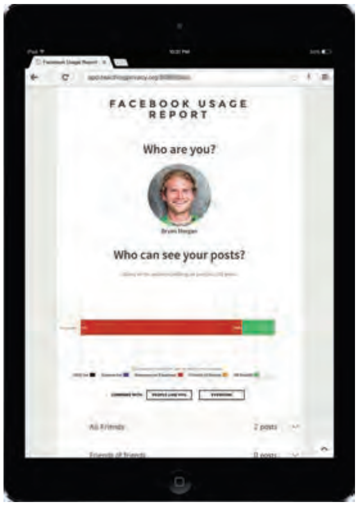
The unusual level of interest from K-12 and college teachers in their multimedia privacy research inspired ICSI's Multimedia Group to begin a small project in which they would work with educators to develop a set of interactive privacy-visualization apps and learning tools for classroom use. These hands-on learning tools were designed to help educators raise students' awareness about the privacy implications of social media use, especially of multimedia sharing. The project quickly grew, attracting an interdisciplinary team of researchers working on privacy.

Our first learning tool, a production-quality app called *Ready or Not?*, was intended to draw students' attention to the risks posed by geotagged social media posts. Given a username, it pulls recent Twitter or Instagram posts from the sites' APIs and uses attached GPS metadata to create a heatmap and timeline of where that user has been posting from recently. It then gives prevention tips for how to keep social media posts from giving away location information (see Figure 1).

In addition to being used in classrooms, Ready or Not? also inspired stories by several high-profile national and local news agencies, drawing the attention of a much larger audience to the risks of geotags.[8–10] As of November 2014, more than 25,000 unique users had tried the app.

To date, the Teaching Privacy project has also created two other interactive apps tailored to help young people visualize how

(a)



(b)

Figure 2. Teaching Privacy online learning tools. (a) Oh, the Places You (and Your Data) Will Go! and (b) Social Media Usage Report.

information persists and travels on the Internet along with a number of classroom activities and two videos. Here is a sampling of these learning tools:

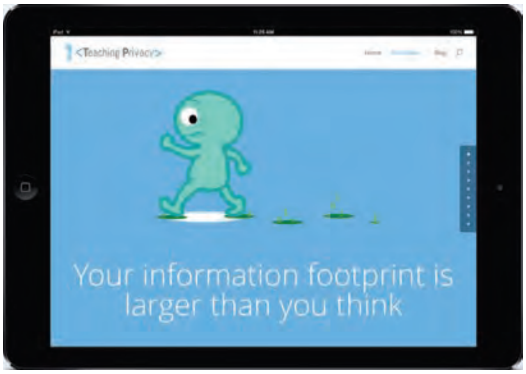
- *Oh the Places You* (*and Your Data*) *Will Go*!: This choose-your-own-adventure activity for the classroom illustrates principles of privacy we deal with in daily life. Students make privacy decisions for a hypothetical person and see what happens. The result is interspersed with teaching material about the implications of those choices and encourages users to actively engage with privacy issues (see Figure 2a).
- *Social Media Usage Report*: This app, which doubles as an evaluation tool, allows a user to visualize how many people see different types of information about them on Facebook and compare themselves with other people who have similar personalities. Along with the visualization, the app gives detailed instructions and advice on changing one's Facebook privacy settings (see Figure 2b).
- *Welcome to the Internet*: Aimed at college or advanced high school level classes, this stand-alone online lab demonstrates how the structure of the Internet affects students' privacy. It includes a technical introduction to concepts like routing, host redirects, and browser signatures as well as exercises based on the Teaching Privacy content and website.
- *Digital Footprints*: The first in our series of classroom-ready educational videos, Digital Footprints explores the many factors and activities that add to each person's ever-growing information footprint and touches on some of the strategies that can limit it. The narration is accompanied by live-drawn humorous visuals that turn each point into a memorable story.

In addition to the authors, Eungchan Kim, Arany Uthayakumar, and Ketrina Yim helped produce the tools described here. All of these resources are accessible via our Teachers' Portal at www.teachingprivacy.org/teachers-portal.

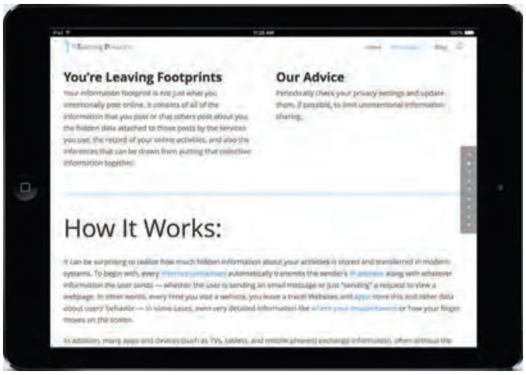
A Curriculum for Teaching Privacy

As we were developing the educational apps, we also began to develop a content base explaining how and why personal information travels around the Internet, along with practical guidance about how young people can better protect themselves online, given the facts on the ground. The project team together identified "Ten Principles for Online Privacy," a set of fundamental, but often counterintuitive, precepts around which to focus the more extensive explanations and suggestions. For example, our first principle, "Your information footprint is larger than you think," draws attention to important (but perhaps esoteric) technical concepts like metadata, device identifiers, and data-mining and inference techniques.

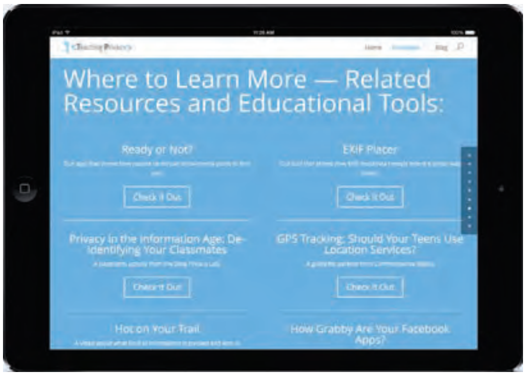
The apps and content base are hosted on the Teaching Privacy website, www.



(a)



(b)



(c)

***Figure 3.** Excerpts from the Teaching Privacy webpage explaining the privacy principle "Your information footprint is larger than you think." (a) Principle, (b) summary, and (c) resources. (Illustration by Ketrina Yim.)*

teachingprivacy.org (see Figure 3). The website includes a page for each of the Ten Principles, with an easy-to-understand description of the underlying technical and social concepts; suggestions for actions people can take to better protect their privacy; "ignite" questions to provoke critical thinking; and links to related resources, learning tools, and guides.

To make the resources we are creating easier to integrate into classroom lessons, we are currently developing a set of flexible, classroom-ready teaching modules and a teachers' guide, together called TROPE (Teachers' Resources for Online Privacy Education). We are making these materials available at http://teachingprivacy.org/teachers-portal as we build them. We invite educators to use them and provide feedback on how to make them most effective as classroom tools across a variety of situations.

As public concern over online privacy in the United States has grown in the last few years,[11] more educational resources have become available. However, we believe Teaching Privacy is unique in our combination of accurate, accessible technical details; comprehensive coverage; and attention to practical strategies.

Despite increased public awareness of privacy issues, most people still do not have a good handle on the specific mechanisms involved, nor the steps they can take to protect their privacy online. In addition, a major challenge we have faced in designing learning tools is how to engage young people in actively thinking about their online privacy without resorting to scare tactics. Such tactics can unintentionally suggest that there's no point in even trying to manage one's online privacy, which is problematic given that, in our interactions with students, we found that many young people are already close to relinquishing the idea of controlling their privacy altogether.

Our approach therefore focuses on linking awareness of the specific privacy implications of social multimedia with knowledge of actionable, practical strategies for managing privacy. For example, in the case of the Ready or Not? app, we teach users about the inferences about daily habits that can be drawn from GPS-tagged images and then show them how they can control which apps use location services. We also encourage them to use those strategies proactively. One of our objectives in Teaching Privacy is for young people to understand that there is always something they can do to manage their online privacy, whether by changing their privacy settings, choosing different services, or communicating with friends and family about their privacy preferences.

Spreading the Word and Gaining Inspiration

Throughout the project, we have been sharing our learning tools and materials through multiple curriculum-resourcing and professional development channels associated with Berkeley's CS 10 course, The Beauty and Joy of Computing (BJC, http://bjc.berkeley.edu). BJC engages non–computer science majors with

Sharing our Curriculum with Teachers and Students, Worldwide

By Daniel D. Garcia

The Teaching Privacy materials have been incorporated at UC Berkeley through our The Beauty and Joy of Computing (BJC) nonmajor course. BJC draws in students at all levels, freshman through graduate (and staff), from every department on campus. The course has been chosen twice as a College Board Advanced Placement Computer Science Principles (AP CSP) pilot course, recognizing it as a model university course that covers the learning objectives in their curriculum framework.

We teach beginning programming using Snap! (http://snap.berkeley.edu/), a friendly, blocks-based language; cover several big ideas like abstraction, recursion, and higher-order functions; and discuss the social implications of computing. The Teaching Privacy materials fit perfectly into that last theme, where we highlight the trade-offs many people make with computing innovations, balancing convenience with the privacy implications.

For example, many smartphone users appreciate automated photo geotagging because it allows their photos to be automatically organized by location. They often don't realize (or do realize and find it worth the trade-off) that sharing geotagged photos on the Web can reveal the location of their home and even tell potential burglars when their home is vacant.

BJC reaches 700 UC Berkeley students per year. Since 2010, the National Science Foundation has provided us with funding so we can offer professional development (PD) to high school teachers for our BJC curriculum. To date, more than 200 high school teachers have been to our summer PD sessions, and many are now teaching this material in their high schools. We have just received another NSF grant to take the course to 100 more teachers in New York City, the nation's largest and most diverse school system. (See the Bringing BJC to New York City High Schools webpage, http://bjc.berkeley.edu/website/bjc4nyc.html, for more details.) We are currently developing a massive open online course (MOOC) version of our BJC course, entitled BJCx, that will launch on Labor Day 2015.

With the AP CSP exam beginning in the spring of 2017, we look forward to this course, and the Teaching Privacy material contained within it, reaching thousands of students across the country and the world.

---

technological concepts, including the social implications of computing (see the "Sharing Our Curriculum" sidebar for more details).

We receive valuable feedback by using the learning tools and materials in BJC. For example, we presented students with an early draft of the Ten Principles with brief explanations and asked which were most surprising, which explanations were confusing, and what we might be missing. This input helped us refine the principles and explanations. For example, many students pointed out that the seeming contradiction between the principles "Identity is not guaranteed on the Internet" and "There is no anonymity on the Internet" could be confusing, so we revised the materials to explain explicitly how both can be true, depending on the resources and technical knowledge of both the person trying to hide their identity and the person trying to figure it out.

We also discuss the resources with high school educators and pilot-test them with BFOIT students, who have helped us identify which approaches are likely to be most engaging. Interns from both BJC and BFOIT have provided guidance on what their peers in our target demographic do and don't know regarding online privacy issues, provided critical feedback on content and learning tools, and even led the creation of some learning tools.

Figure 4. Teaching Privacy demonstration at the Cal Day open house in April 2014. UC Berkeley student Madeeha Ghori demonstrates the Ready or Not? interactive app and explains the implications to prospective Berkeley students and their parents. (Photo by Bryan Morgan.)

In more general outreach, we have presented Teaching Privacy in public lectures and in a popular "What Does the Internet Know

> **Experts can provide the public with a realistic understanding of what individuals actually can and cannot do to protect their privacy, in practical terms.**

About You?" interactive lab attended by hundreds of high school students and their parents during an open house at UC Berkeley (see Figure 4).

We will be introducing the TROPE materials in March 2015 at the annual gathering of the ACM's Special Interest Group on Computer Science Education (SIGCSE). As well as introducing the materials, this workshop will allow us to solicit feedback and on-the-ground stories, so we can gain a better understanding of specific problems faced by both teachers and students.

Engaging with people about online privacy at these public events not only improves and gains buy-in for the Teaching Privacy project, it also provides new inspiration and provokes new questions for our respective research programs.

Leveraging Expertise to Make a Difference

The current lively public discussion about online privacy provides a new opportunity for technical experts to contribute to an informed dialogue. In particular, experts can provide the public with a realistic understanding of what individuals actually can and cannot do to protect their privacy, in practical terms. The Teaching Privacy project provides one model for such contributions on a national scale (see the "Sharing Our Curriculum" sidebar), with multimedia researchers contextualizing practical advice in explanations of how multimedia content creation and distribution work in the context of Internet architecture.

As well as demonstrating the utility of contributions from multimedia experts, Teaching Privacy has shown the importance of cross-disciplinary collaboration in making those contributions most effective. The expertise of educators and social scientists—for example, Gordo's fieldwork with community college students and parents of school-age children—has grounded our work in a broad understanding of what different sectors of the public know (or don't know) about online privacy and how best to reach them. This understanding increases our potential to engage people about actively managing their privacy. In addition, interaction with computer science education also increases our potential to spark the interest of young people in developing and researching (not just using) multimedia technology.

We hope that the successes of this project so far will raise awareness among computer scientists and engineers like the readers of this magazine, not only of the essential need to consider privacy concerns in researching and designing multimedia applications, but also of the potential active contributions we can make by reaching out to nonexperts and supporting public understanding of the ideas and principles that engage us most.

Lastly, we are asking you, the readers, to help with your expertise. Contact us and tell us your privacy story or about your privacy-related multimedia application. We hope to link more and more interesting projects and stories from our website, making it an ever-growing resource that anyone can use as a basis for teaching or learning about privacy. **MM**

Acknowledgments

Teaching Privacy has been the work of a whole team of researchers, educators, and interns. In addition to the authors listed here, project contributors include Alexis Conway, Orpheus Crutchfield, Isha Doshi, Melia Henderson, Jeffrey Jacinto, Eungchan Kim, Itzel Martinez, Gerardo Sánchez, Robin Sommer, Arany Uthayakumar, and Ketrina Yim.

Teaching Privacy is supported in part by National Science Foundation grants CNS-1065240: Understanding and Managing the Impact of Global Inference on Online Privacy and DGE-1419319: Teachers' Resources for Online Privacy Education as well as by the National Telecommunications and Information Administration's Broadband Technology Opportunities Program, through the California

Connects program administered by the Foundation for California Community Colleges. Any opinions, findings, and conclusions expressed in this article are those of the authors and do not necessarily reflect the views of the funders.

References

1. H. Lei et al., "User Verification: Matching the Uploaders of Videos across Accounts," *Proc. IEEE Int'l Conf. Acoustic, Speech, and Signal Processing* (ICASSP), 2011, pp. 2404–2407.
2. X. Anguera Miro et al., "Speaker Diarization: A Review of Recent Research," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 20, no. 2, 2012, pp. 356–370.
3. G. Friedland, O. Vinyals, and T. Darrell, "Multimodal Location Estimation," *Proc. ACM Int'l Conf. Multimedia*, 2010, pp. 1245–1251.
4. M. Larson et al., "Automatic Tagging and Geotagging in Video Collections and Communities," *Proc. 1st ACM Int'l Conf. Multimedia Retrieval (ICMR)*, 2011, article no. 51.
5. G. Friedland and R. Sommer, "Cybercasing the Joint: On the Privacy Implications of Geotagging," *Proc. 5th USENIX Workshop on Hot Topics in Security (HotSec)*, 2010, pp. 1–8.
6. B. Gordo, "Developing a Framework for Understanding Online Privacy. Appendix K," *California Connects: Improving Digital Opportunities in Underserved California Communities*, Sept. 2013. http://digitalequality.net/understanding-online-privacy.
7. O.S.L. Crutchfield et al., "Berkeley Foundation for Opportunities in Information Technology: A Decade of Broadening Participation," *ACM Trans. Computing Education*, vol. 11, no. 3, 2011, article no. 15.
8. A. Truong, "Privacy App Pinpoints Your Exact Location Using Social Media," *Fast Company*, 3 Sept. 2013; www.fastcompany.com/3016662/fast-feed/privacy-app-pinpoints-your-exact-location-using-social-media.
9. K. Shubber, "Mapping Websites Reveal Just How Stupid It Is to Geotag Your Tweets," *Wired UK*, 4 Sept. 2013; www.wired.co.uk/news/archive/2013-09/04/twitter-geotagging.
10. J. Watts, "ConsumerWatch: Social Media Users May Be Revealing Too Much about Location," video, KPIX/CBS SF Bay Area, 4 Nov. 2013; http://sanfrancisco.cbslocal.com/video/9491297-consumerwatch-social-media-users-may-be-revealing-too-much-about-location/.
11. M. Madden, "Public Perceptions of Privacy and Security in the Post-Snowden Era," Pew Research Center, Nov. 2014; www.pewinternet.org/2014/11/12/public-privacy-perceptions/.

Julia Bernd is a linguist and social science researcher who is currently adding interdisciplinarity to the Audio & Multimedia Group at the International Computer Science Institute, a nonprofit research institute affiliated with the EECS Department at the University of California, Berkeley. Contact her at jbernd@icsi.berkeley.edu.

Blanca Gordo is a social scientist and senior researcher in the Artificial Intelligence Group at ICSI and a visiting scholar at the Institute for the Study of Societal Issues at UC Berkeley. Contact her at blanca@icsi.berkeley.edu.

Jaeyoung Choi is a computer scientist and staff researcher at ICSI, working on multimodal location estimation of videos. Contact him at jaeyoung@icsi.berkeley.edu.

Bryan Morgan is a multimedia designer who works on Teaching Privacy at ICSI as well as a master's student in information management and systems at UC Berkeley. Contact him at brynamo@icsi.berkeley.edu.

Nicholas Henderson is the program director for Science for Youth (SCI-FY), the middle school component of the BFOIT computing outreach program at ICSI.

Serge Egelman is a senior researcher in the Networking & Security Group at ICSI, and he holds an appointment in UC Berkeley's Electrical Engineering and Computer Sciences Department. Contact him at egelman@icsi.berkeley.edu.

Daniel D. Garcia is a senior lecturer SOE in the EECS Department at the University of California, Berkeley, the codeveloper of the BJC CS Principles curriculum, and an ACM Distinguished Educator. Contact him at ddgarcia@cs.berkeley.edu.

Gerald Friedland is the director of the Audio and Multimedia Group at ICSI and a part-time lecturer in the UC Berkeley EECS Department. Contact him at fractor@icsi.berkeley.edu.

This article originally appeared in IEEE MultiMedia, *vol. 22, no. 1, 2015.*

Editor: Yih-Farn Robin Chen • chen@research.att.com

Mobile Video Delivery: Challenges and Opportunities

Rajesh Krishna Panta • *AT&T Labs Research*

Demand for mobile videos is significantly on the rise, yet network infrastructure resources that deliver mobile videos haven't increased proportionally. In exploring the challenges and solutions for mobile video delivery, the author also considers the impact of virtualizing network functions and services.

Video constitutes a large fraction of the Internet traffic today. Globally, IP video traffic will be 79 percent of all Internet traffic in 2018, excluding video shared through peer-to-peer file sharing.[1] In the last few years, we've seen a tremendous increase in cellular data traffic, fueled primarily by the surge in mobile video traffic. We can mostly attribute the tremendous growth of mobile video traffic to the popularity of smart handheld devices like smartphones and tablets, and the ease with which we can generate and consume videos using mobile devices.

This ever-increasing demand has introduced several challenges in delivering videos with a good quality of experience to the user. Although these challenges exist in both wired and wireless environments, the problem is especially severe in cellular mobile contexts due to several factors: for example, the limited wireless spectrum and bandwidth, time- and location-dependent wireless link characteristics, radio congestion, potential handoff issues, heterogeneous device features and limitations, and so on. Although network operators have made huge investments in recent years to address this problem, the increase in network resources hasn't kept pace with the surge in mobile traffic demand. As a result, the quality of the user's experience for mobile video is of critical concern today.

With this in mind, here I explore several techniques for efficient mobile video delivery. I also consider why current systems might be inadequate in the long run. Finally, I discuss the recent trend of virtualizing mobility core (the network that connects the radio segment of the cellular network to the IP network), which can significantly impact mobile video delivery. The focus of this article is on mobile video streaming, not video downloads, because most of the popular video services like YouTube and Netflix use streaming.

Mobile Video Streaming Background

Existing popular video streaming services use an HTTP-based over-the-top method to deliver video content to users. Early versions of such services pushed video content to the client devices as quickly as possible using TCP at a constant bit rate, irrespective of network conditions and device capabilities. However, network conditions can vary greatly for different clients. Even for a single client, the available network bandwidth can vary over time. Furthermore, because of heterogeneous device capabilities, the quality of video a client can play varies significantly among client devices.

These problems are even more pronounced in wireless cellular environments because of several factors, including dynamic radio environment, device mobility, and a wide range of device capabilities. Thus, most of the streaming services today support streaming videos at variable bit rates using various adaptive bit-rate (ABR) technologies such as Dynamic Adaptive Streaming over HTTP (DASH), Apple HTTP Live Streaming (HLS), Microsoft Individualized-Integrated Book (IIB) Smooth Streaming, and Adobe HTTP Dynamic Streaming. In these ABR schemes, the source video is divided into smaller chunks and multiple versions of each chunk are pre-encoded at different bit rates on the streaming server. The video player running on the client device adaptively selects the appropriate bit rate based on current network conditions and device capabilities. The client uses a prediction algorithm to estimate the network bandwidth that will be available to download the next video chunk, based on the observed TCP throughput for previous video chunks. Each video chunk is typically stored as a regular file on the video server and downloaded using standard HTTP GET requests. In short, the ABR mechanism's goal is to maximize the video quality by choosing the highest bit rate the network can support without causing video pauses for the client.

If network conditions are stable, the device can make fairly accurate estimates of future capacity based on past TCP throughput observations. However, in wireless cellular environments, network conditions can vary widely over time. Mobility complicates the problem further, since available network capacity can change significantly as a function of the user location — if a user moves away from the base station, the available bandwidth can degrade significantly. Several studies show that the throughput observed during a single video streaming session can vary wildly. For example, even within a single Netflix session, the measured throughput varies from 500 Kbits/s to 17 Mbits/s.[2] Thus, estimating future network capacity is challenging in mobile video streaming. Inaccurate estimates can lead to degraded quality of experience for the end users. If network capacity is underestimated, the user will receive the video with lower quality, even though the current network condition allows a higher quality of video to be delivered to the user. If the ABR mechanism overestimates the network capacity, the video player picks a video bit rate greater than network capacity. As a result, the video plays back faster than the rate at which it's downloaded. This ultimately leads to video buffer depletion, causing video pauses.

Challenges and Opportunities

Now that we have this background, let's discuss some important challenges for efficient mobile video delivery. We consider some potential solutions as well.

Rate Adaptation and Scheduling

As previously mentioned, the fundamental problem in ABR-based mobile video streaming is the inaccurate estimation of network bandwidth. There are many factors that make network bandwidth estimation challenging in mobile networks, such as the time and location varying network link conditions, device mobility, congestion in the wireless medium, random device arrival/departure, and so on. The most commonly used methods for estimating network bandwidth are based on the observed TCP throughput for downloading previous video chunks. Some recent research studies[2] propose the use of other parameters, such as the video player's buffer occupancy at the client as an indicator of network conditions. The idea is based on the assumption that the rate at which the video buffer is filled represents approximately the current available network bandwidth. The fundamental problem with these approaches is that in wireless environments, the available network capacity can vary so wildly that any prediction based on historical information might not always yield accurate bandwidth estimation.

Another fundamental problem of ABR-based video streaming is that when a cellular network link is shared by multiple ABR flows, the temporal overlap of the chunks of different flows might cause under- or over-estimation of the network bandwidth.[3] In cellular networks, the base station schedules flows of different users to achieve some notion of fair resource allocation across end user devices in a cell. However, these schedulers are designed for single bit-rate videos and they can't handle multiple bit rates efficiently. As a result, sharing a cellular network link among multiple ABR flows of varying bit rates results in inefficient use of radio resources, unfair resource allocation among concurrent video streaming flows, and instability of bit-rate selection. By separating the scheduling algorithm of ABR video flows from other non-ABR video flows (and non-video flows), and by designing the scheduler optimized for multiple ABR flows, we can achieve better resource utilization and fair scheduling among such flows.[4]

This discussion suggests that different schedulers customized for different types of data flows could be necessary to ensure better performance for different classes of data traffic. It's also reasonable to assume that various types of information available at the network might help solve (or mitigate) several of the aforementioned problems for mobile video streaming. For example, in current approaches, mobile clients use local information (such as TCP throughput and buffer occupancy) alone to make future bandwidth estimation. However, if the information available in the cellular network (for example, congestion level in a cell or the total number of active ABR flows in cell) is somehow communicated to mobile devices, they might be able to make better bandwidth estimates. If the network provides some APIs to share this information with video streaming services (and other applications), then those services are equipped better to tackle many of mobile video streaming's fundamental problems. Furthermore, each video service might decide to implement its own customized scheduler, policy engines, and data-processing algorithms, based on their requirements and capabilities. Later, we discuss why this is difficult in today's environment and how the trend towards virtualizing mobility core can help in achieving these goals.

Demand-Supply Mismatch

Another critical issue in mobile video delivery is the mismatch between mobile video demand and the network resources available to meet such demands. Although mobile video demand has been increasing tremendously in recent years, the capacity of cellular networks, especially the wireless spectrum, hasn't increased proportionally. One way to tackle this problem is to increase network resources by, for example, using small cells to augment the capacity of traditional macro cells, adding WiFi hotspots to offload cellular traffic to WiFi, and using portable base stations (such as Cell On Wheels or COWs) in crowded areas for some time periods. These approaches are expensive, though, and they might be insufficient to handle the explosive growth of video traffic.

My colleagues and I came up with a recent proposal, CoAST,[4] that attempts to solve this problem by using two key insights. First, the mobile data traffic exhibits high burstiness over small time scales (tens of seconds). Thus, to ensure adequate quality of service at all times, it's important to reduce instantaneous peak traffic, not just aggregate traffic. Second, although it might not seem intuitive at first glance, the video streaming clients can tolerate delays of tens of seconds (as long as the playback buffer isn't empty) without affecting the quality of video playback. These two insights suggest that if the right video flow (from the set of all video flows in a cell) is delayed at the right time for the right time duration, it's possible to reduce peak traffic in a cell without affecting the user experience on any mobile device. However, this requires both device-level information (for example, delay tolerance values at the given time instant) and cell-level information (for instance, the total traffic demand in the cell at a given time instant). CoAST provides an efficient collaboration framework for mobile devices and a cellular network to exchange such information and make proper decisions about delaying video streaming traffic. CoAST reduces traffic peaks in a cell by up to 50 percent and eliminates some pauses in video streaming.

Our previous work[5] shows another example of how the collaboration between mobile devices and cellular infrastructure can help improve the performance of video streaming. Cellular networks possess rich information about mobile devices' location, mobility patterns, and possible geographic regions where network link qualities might be degraded. By analyzing this information, the network can predict and inform a mobile device running a video streaming application that it's likely to enter a congested region in the near future. Upon receiving such notification, the mobile device can aggressively download the video content. As a result, when the mobile device actually enters the congested region, it will have sufficient video content in its playback buffer to avoid pauses.

These two examples highlight the importance of cooperation between mobile devices and network infrastructure to improve the performance of video streaming in cellular networks. And of course, there are many other examples that could benefit from such collaboration. As we discuss later, the trend of virtualizing network functions can facilitate the development of such collaborative frameworks.

Multicast Videos

Multicast is another way to improve the performance of mobile video delivery in some cases. If a group of users in a given area are interested in the same video at the same time, it's more efficient in terms of resource usage to broadcast/multicast a single (or a limited number of) video flow to the all the users, instead of transmitting individual unicast streams to each mobile device. Enhanced Multimedia Broadcast Multicast Service (eMBMS) is the 3GPP (Third Generation Partnership Project) standard specification for multicast over Long-Term Evolution (LTE) networks. The multicast feature is especially useful in scenarios such as live video streaming for broadcasting sporting events, concerts, or other special events. There are several technical challenges for proper implementation of the multicast service. For example, how should we allocate network resources among unicast and multicast services? Because different users in a cell can experience different radio link conditions, how should we organize users in multicast groups so that mobile devices can receive video with quality commensurate with their wireless link bandwidth? These issues suggest that we might need customized schedulers and distribution algorithms for multicast videos, further strengthening the case for virtualized mobility core.

The Effect of Network Virtualization

In recent years, there's a trend in the telecommunication industry towards an important transformation — virtualizing network functions and services. The idea is to replace the dedicated network appliances that perform specific network functions with general-purpose commodity hardware (such as x86-based machines), where the network functions are implemented in software. The advantages of such a transformation include a reduction in capital and operational costs, shortening the time required to deploy new networking services to support changing business requirements, and the ability to scale up or down services to address changing demands.

This trend has important implications for mobile video streaming as well. From the discussion in previous sections, it's clear that improving video delivery over cellular networks might require customized solutions for different scenarios. We might need highly customized schedulers and policy decisions for different video services. The collaboration mechanism between mobile devices and the cellular network might be different for different video streaming services (and other non-video applications). Multicast solutions may require yet another set of resource allocation modules. It's difficult to incorporate these highly individualized solutions into today's appliance-based network elements, which afford little flexibility.

Virtualizing network functions enables the development of a software platform that multiple video streaming (or non-video) applications can share, and then the applications can implement their own high-level network functions according to their requirements and capabilities. In addition to the ability to deploy customized solutions for individual streaming services, the virtualized mobility core also allows allocation of resources for each application based on their current load. Efficient resource utilization is important for network service providers because it helps to tackle the mismatch between high video demand and limited resources available to satisfy such demands. For example, a video streaming service can independently scale out its network resources when it experiences a burst in the mobile video demand. Although this subsection focuses on how virtualization can help video delivery in mobile networks, we should note that many other applications could also benefit from the virtualized platform.

One important question regarding network function virtualization is: Can network functions implemented in software meet critical performance requirements, such as latency and throughput? It turns out that recent technological developments like Intel's Data Plane Development Kit (DPDK) help narrow the performance gap between hardware and software implementations.

Even though we've come to expect high-quality videos on mobile devices, the increased demand for mobile videos means that creating such an experience for users is challenging. Hopefully some of the techniques mentioned here (customized scheduling at the network edge, and collaboration between the network and mobile devices) can diminish these challenges. The important features of network function virtualization also show potential in providing a software platform to build resource-aware and customized solutions for mobile video streaming services.

References

1. Cisco, *Cisco Visual Networking Index: Forecast and Methodology, 2013–2018*, white paper, 10 June 2014; www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.pdf.
2. T.-Y. Huang et al., "A Buffer-Based Approach to Rate Adaptation: Evidence from a Large Video Streaming Service," *Proc. ACM Special Interest Group on Data Comm.*, 2014, pp. 187–198.
3. J. Chen et al., "A Scheduling Framework for Adaptive Video Delivery over Cellular Networks," *Proc. ACM Ann. Int'l Conf. Mobile Computing and Networking*, 2013, pp. 389–400.
4. C. Shi et al., "CoAST: Collaborative Application-Aware Scheduling of Last-Mile Cellular Traffic," *Proc. ACM Int'l Conf. Mobile Systems, Applications, and Services*, 2014, pp. 245–258.
5. N. Theera-Ampornpunt et al., "Using Big Data for More Dependability: A Cellular Network Tale," *Proc. ACM Workshop on Hot Topics in Dependable Systems*, 2013, article no. 2.

Rajesh Krishna Panta is a member of the technical staff at AT&T Labs Research. His research interests include cloud computing, Big Data, storage systems, mobile systems, wireless networks, and wireless sensor networks. Panta has a PhD in electrical and computer engineering from Purdue University. Contact him at rpanta@research.att.com.

This article originally appeared in IEEE Internet Computing, *vol. 19, no. 3, 2015.*

Susanne Boll
University of Oldenburg

Photos to Remember, Photos to Forget

Wolfgang Nejdl and Claudia Niederée
L3S Research Center and Leibniz Universität Hannover

Would you like to remember (almost) everything? If yes, you might reconsider after reading the next few stories.

Have you ever heard of Ireneo Funes? He is the main character in "Funes the Memorious," a short story by the Argentinian writer Jorge Luis Borges, first published in 1942.[1] Ireneo can remember everything, but it turns out that this bogs him down, and he becomes lost in details. He spends his time investigating ideas like inventing a new system of enumeration, which gives each number a unique name (such as "Maximo Perez" instead of "seven-thousand-and-fourteen"), or enumerating and cataloging all his memories in a day. Coping with new details every second makes him unable to focus and to generalize, and he spends his short life lost in thought.

This phenomenon also happens in real life. Jill Price is a California woman who through a rare combination of genes remembers virtually everything that has happened to her since the age of 15.[2] "Starting on Feb. 5, 1980, I remember everything. That was a Tuesday."[3] She also remembers the sensations that came with each event, which makes looking back painful at times. Thus, remembering dominates her life.

"The Web remembers everything" is another phrase that is used a lot these days. While this is not true—Web archiving in reality is spotty at best—things do turn up on the Web that were not meant to be remembered. In addition to the two stories already cited, Viktor Mayer-Schönberger has discussed the "Drunken Pirate" case of Stacy Snyder, a teacher in training and student at the Millersville University School of Education.[4] In 2006, Snyder posted a photo on MySpace with the caption "Drunken Pirate" that showed her drinking at a private party wearing a pirate head. This photo subsequently lead to her being denied a teaching degree, with the dean saying that she was promoting drinking in "virtual view" of her underage students. You can still find the photo online.

Prompted by such cases, the "right to be forgotten" has been a theme of the European Commission in recent years. We invite the interested reader to read the Snyder case[4] as well as an interesting discussion printed as a special symposium issue of the *Stanford Law Review* online on "The Privacy Paradox."[5] A new presentation by the European Commission on the "right to be forgotten" ruling of the EU Court on 13 May 2014 is informative as well; the presentation discusses the scope of this ruling and its relationship to the proposed EU Data Protection Regulation.[6]

In this article, we mainly focus on this question: What should we remember and thus archive, and what can we forget? Each year makes it easier to accumulate large numbers of photos and videos in the social and personal digital space. Their long-term existence is mostly driven by chance rather than by clear guidelines or rules for archiving them. Thus, unfortunately, cases of nonintended disappearance of personal photos happen much too often, caused by a similar lack of survival control as in the cases of extensive survival of "youthful" photos in the public space. How should we proceed? Is it really as easy as providing a "date to forget" to any digital information for an improved survival control, as Mayer-Schönberger[4] speculatively has suggested?

Managed Forgetting and Contextualized Remembering

The European project ForgetIT (www.forgetIT-project.eu) is investigating the introduction of a form of digital or managed forgetting into information management environments. The project, managed by the L3S Research Center at the Leibniz Universität Hannover, focuses on the idea of making more conscious decisions about which content is really important, and

 2376-113X/15/$31.00

thus should be preserved safely, and which content we can (and should) forget. Although we often follow a keep-it-all strategy for our content, in reality the content is subject to a form of random forgetting, for example, as a result of hardware crashes or when storage formats and technologies become obsolete. What we need instead is a principled strategy of forgetting, preservation, and remembering.

Inspired by the central role human forgetting plays in helping us stay focused, in ForgetIT, *managed forgetting* supports resource selection for preservation. Preservation refers to technological and organizational methods, which ensure a long-term survival and accessibility of content over several decades and longer. At first glance, forgetting and preservation might seem to contradict one another.[7] However, because preservation comes at a cost, it is important to make conscious decisions about which resources to invest in. In ForgetIT, managed forgetting is complemented by *synergetic preservation,* which makes preservation technology easier to use in personal information management. Finally, the third cornerstone of the project is *contextualized remembering,* which equips preserved content with sufficient and evolving context information. Such a managed approach will keep our archives understandable and useful, thus ensuring a type of semantic survival of the selected resources. This is inspired by the constructive nature of the human remembering process. Personal preservation—how we deal with our personal information such as photo collections over time—is one of the application domains of the ForgetIT project.

Which Photos Should We Archive?

In a first systematic research work on automated photo selection for preservation,[8] we asked 35 users to select 20 percent from one of their own collection of photos using the metaphor of a digital vault (see Figure 1). This analysis of untagged photo collections of up to several hundreds of photos per collection and more than 8,000 images overall was used to analyze the factors (and features) that dominate such photo selection decisions in real-world settings—for example, how might you select photos from a collection brought back from a holiday trip? The photo selection task was accompanied by a user survey, which elicited attitudes towards preservation and perceived photo selection criteria.

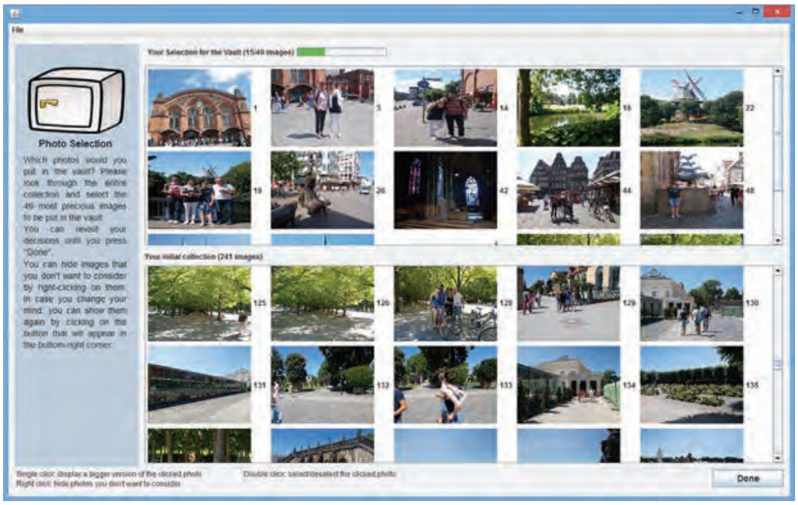


Figure 1. Protected by digital vaults. The ForgetIT project asked users to select 20 percent of their personal photo collections for preservation to analyze the factors (and features) that dominate photo selection decisions.

The captured selection decisions of the users were employed for learning the most dominant features for automated photo selection methods. To develop these methods, two observed behaviors of users in photo selection were considered (see Figure 2): First, users tended to follow a *reduction-oriented strategy* of getting rid of duplicates and near duplicates to gradually reduce the decision space. This behavior inspires the development of reduction-oriented methods for automated photo selection. Second, users tried to cover different subevents with the set of photos that they selected. This motivates the development of *coverage-oriented methods,* aiming to cover all the subevents of a holiday trip, for example. Event coverage was also often named as an important selection criterion in the accompanying survey. It has to be noted, however, that subjective measures such as that a photo “evokes a memory” or “is important to me” were rated even higher as reasons for photo selection by the users in the survey. This suggests that automated photo selection is a difficult task driven by a variety of partially subjective factors.

Starting from the observed reduction-oriented and coverage-oriented user behavior, we developed and compared coverage-based and reduction-oriented methods for photo selection.[8] The enabling mechanism underlying both approaches is *photo importance prediction,* which is learned based on a rich set of features including quality-based, face-based, and collection-level features. A special focus was

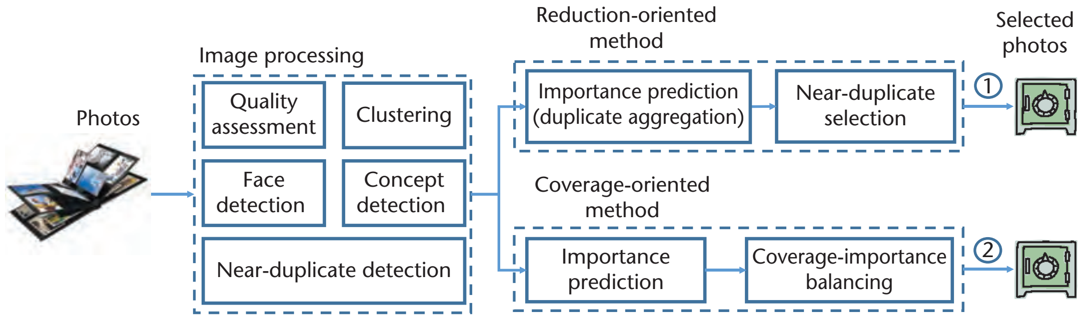


Figure 2. Automated photo selection for preservation. Following user observed behaviors, the developed methods use reduction-oriented strategy to prune collections and coverage-oriented methods to ensure all subevents are represented.

placed on concept-based features, which rely on a method for detecting high-level concepts in a photo in a two-step process,[9] thus at least partially capturing the semantics of a photo. For the coverage-oriented approach, this is combined with a semantic clustering method,[10] which relies on concept detection and subsequently clusters photos based on those concepts and time, aiming for an event-based grouping of the photos. Importance prediction for the individual photos is used to select photos from the individual clusters. For the reduction-based approach, duplicates are aggregated based on a relaxed near-duplicate detection method, and subsequent photo selection (based on importance prediction) considers them as a single photo.

The experimental evaluation of the two methods inspired by human behavior shows that both the reduction- and coverage-based approaches are promising solutions to address the difficult task of selecting photos for preservation, with the former outperforming the latter. In analyzing the relevant features, we clearly observed that the use of semantic information (concepts and faces in image) and collection-based features are core ingredients for tackling this problem. Our experiments also confirmed the user survey results, suggesting that image quality is only a secondary criterion for this photo selection task.

Other Work on Photo Selection

Automated photo selection has been studied in a variety of settings including photo summarization,[11,12] selection of representative photos,[13,14] and automated creation of photobooks[15] and collages.[16] A variety of factors influences the selection of photos from a photo collection. Clearly, the mix of factors depends on the photo selection purpose. For example, we would use different criteria to pick photos for sharing on Facebook than to decide which photos to preserve for the long term.

The development of photo selection methods also has to take into account photo taking and management practices. Early work[17] showed a change in such practices implied by the growing proliferation of digital photography (such as the large increase in the number of photos). Relevant social practices surrounding photos are photo sharing practices driven by the digital medium and social media, the increase in more mundane motives for taking photos (such as food), and the mix of purposes for photo taking (and preservation) including self-presentation, social aspects, remembering, and reflection. (Previous research provides a good overview of such social practices.[18])

Features that have been investigated in the context of automated photo selection are extracted from the content itself (visual features), its context (temporal information), or photo annotations. They include more objective factors such as photo quality, aesthetics, and depicted entities, especially people. Congcong Li and his colleagues presented a method for translating important criteria for photo aesthetics into methods that can be applied for automated photo selection.[19] However, photo selection is also driven by more subjective factors, for example, when a picture captures an important moment or when it displays what the user wanted to capture. A qualitative study on "keep or delete" decisions for photos showed the importance of subjective criteria in photo selection.[20] In this study with 74 participants, three categories were identified as the most frequent reasons for deleting photos: reaction (including photos that have a negative

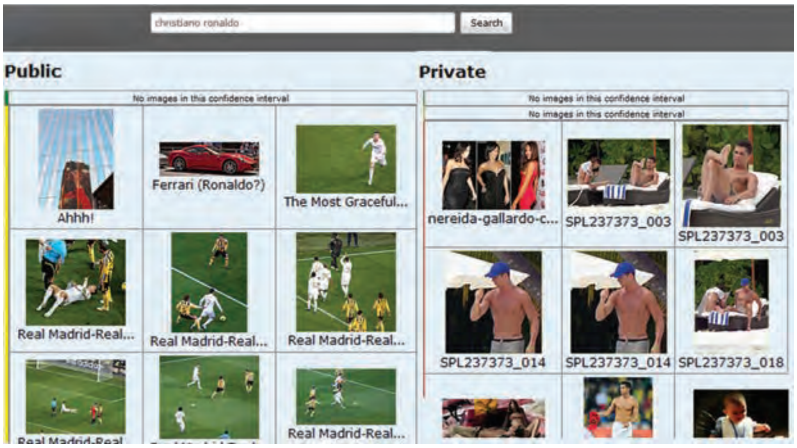


Figure 3. Public versus private photos. The proposed automatic privacy classification method attempts to reduce the risk of unintentional exposure of private photos.

connotation for the user or are not sufficiently interesting to the user), aesthetics (including quality), and no show (the photo does not show what it is supposed to show). The importance of subjective factors was also confirmed by another study,[21] where the inclusion of features learned from the eye-tracking behavior (as a signal of personal interest) of a user into the photo selection task considerably improved results.

In addition, photo selection does not consider individual photos in isolation. Therefore, collection-based features such as coverage are also considered in photo selection approaches,[11,12,15] thus reflecting human behavior in photo selection. For example, one photo summarization approach used a multigoal optimization approach for combining coverage of the concept space defined by the original collection and diversity of the selected photo collection.[12] An analysis on the collection level can also uncover signals for increased user engagement while photo taking. This might for example result in a higher number of near duplicates[13] or in shorter time interval between subsequent shots.[16] Such signals can be important evidence for deciding which photos or photo groups might be more important to the user.

Remembering the Right Things

To close the loop of forgetting and remembering in a digital memory, it is also necessary to consider what is remembered and how "memories" are recalled. One cultural probe study showed that memories are not restricted to special events; people also like to remember more mundane and everyday events of the past because they are regularly repeated, of social value, or exemplary for a person's character.[22] One of the further identified reasons for remembering past things is temporal contrast, such as things that were done differently in the past. Such differences in habits and situations also impose challenges in bringing back digital "memories."

For this case, work on contextualization suggests an approach for adding context information in a time-aware way for easing interpretation of older content (such as after decades).[23] Furthermore, it is also important for digital remembering to find relevant memories in the large amount of stored digital memories.

We investigate a related topic as part of our ERC grant project Alexandria (http://alexandria-project.eu), where we are interested in finding the most relevant results in large-scale, multiversion, noisy document repositories, as we have with Web archives. One approach we evaluated in the Alexandria project was to consider historical query intent, investigating for each aspect of a given query, what time point was this aspect most important and preferred over other time points? This helps, for instance, to "remember" the history of a politician or a controversial topic as reflected in an archive, in an adequate and diverse way.

When Not to Share Photos

As mentioned in the introduction, photo sharing and privacy problems often arise when photos are visible outside the originally intended scope. These might be social scopes, such as private versus work environments or friends versus family environments, or a temporal scope such as teenage photos in an employment environment.

To address this problem, we developed a machine-learning-based method for reducing the risk of unintentional exposure of private photos (see Figure 3).[24] Based on an average community notion of privacy elicited from a social game, we designed an automatic privacy classification method. Our experiments show that by exploiting a set of high-level and object-centered features such as the occurrence of faces, edge-direction coherence, and SIFT (scale-invariant feature transform) features, considerable classification performance can be achieved, even without taking into account textual features.

Photo sharing and privacy problems often arise when photos are visible outside the originally intended scope.

Building the Right Code

As a final remark, let us refer back to Lawrence Lessig's "Code and Other Laws of Cyberspace," first published in 1999, with an updated version in 2006.[25] In this book, Lessig brilliantly reminded us that protection and regulation of privacy (and other values) is not done only by law, but also by our norms, markets, and obviously the architecture of the systems we build as computer scientists. Let us build the right code to manage forgetting and remembering on the Internet and the Web. Let us build code to help protect privacy in the right way! **MM**

References

1. J. Luis Borges. "Funes el memorioso" [Funes the Memorious], *Finzioni*, Adelphi Edizioni, 2014.
2. E.S. Parker, L. Cahill, and J.L. McGaugh,"A Case of Unusual Autobiographical Remembering," *Neurocase*, vol. 12, no. 1, 2006, pp. 35–49.
3. S. Shafy, "An Infinite Loop in the Brain," *The Science of Memory.* Spiegel online, 21 Nov. 2008.
4. V. Mayer-Schönberger. *Delete: The Virtue of Forgetting in the Digital Age*, Princeton Univ. Press. 2009.
5. "The Privacy Paradox: Privacy and Its Conflicting Values," *Stanford Law Review Online, symp. issue*, 2 Feb. 2012 – 12 Apr. 2012.
6. European Commission, "Factsheet on the 'Right to be Forgotten' Ruling," 2014; http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_en.pdf.
7. N. Kanhabua, C. Niederée, and W. Siberski, "Towards Concise Preservation by Managed Forgetting: Research Issues and Case Study," *Proc. 10th Int'l Conf. on of Digital Objects* (iPres 2013), 2013.
8. A. Ceroni et al., "To Keep or Not to Keep: Human Inspired Photo Selection Methods for Personal Preservation," tech. report, ForgetIT, 2014.
9. F. Markatopoulou, V. Mezaris, and I. Kompatsiaris, "A Comparative Study on the Use of Multi-label Classification Techniques for Concept-Based Video Indexing and Annotation," *Proc. 20th Anniversary Int'l Conf.* (MMM 2014), Part I, pp. 1–12.
10. C. Papagiannopoulou and V. Mezaris, "Concept-based image clustering and summarization of event-related image collections," *Proc. 1st ACM Workshop on Human Centered Event Understanding from Multimedia* (HuEvent14) at ACM Multimedia (MM 14), 2014.
11. B.-S. Seah, S.S. Bhowmick, and A. Sun, "PRISM: Concept-Preserving Social Image Search Results Summarization," *Proc. 37th Int'l ACM SIGIR Conf. Research and Development in Information Retrieval* (SIGIR 14), 2014, pp. 737–746.
12. P. Sinha, S. Mehrotra, and R. Jain, "Summarization of Personal Photologs Using Multidimensional Content and Context," *Proc. 1st ACM Int'l Conf. Multimedia Retrieval* (ICMR 11), 2011, pp.
13. W.-T. Chu and C.-H. Lin, "Automatic selection of representative photo and smart thumbnailing using near-duplicate detection," *Proc.* 16th *ACM Int'l Conf. Multimedia* (MM 08), 2008, pp. 829–832.
14. E. Guldogan, J. Kangas, and M. Gabbouj, "Personalized Representative Image Selection for Shared Photo Albums," *Proc. 2013 Int'l Conf. Computer Applications Technology* (ICCAT), 2013, pp. 1–4.
15. M. Rabbath, P. Sandhaus, and S. Boll, "Automatic creation of photo books from stories in social media," *Proc. 2nd ACM SIGMM Workshop on Social Media* (WSM 10), 2010, pp.
16. J. Xiao et al., "Mixed-Initiative Photo Collage Authoring," *Proc. 16th ACM Int'l Conf. Multimedia* (MM 08), 2008, pp. 509–518.
17. K. Rodden and K.R. Wood, "How Do People Manage their Digital Photographs?" *Proc. SIGCHI Conf. Human Factors in Computing Systems* (CHI 03), 2003, pp. 409–416.
18. S.E. Lindley et al., "Collocated social practices surrounding photos," *Int'l J. Human-Computer Studies*, vol. 67, no. 12, 2009, pp. 995–1004.
19. C. Li, A.C. Loui, and T. Chen. "Towards Aesthetics: A Photo Quality Assessment and Photo Selection System," *Proc. ACM Int'l Conf. Multimedia* (MM 10), 2010, pp. 827–830.
20. M.K. Wolters, E. Niven, and R.H. Logie, "The Art of Deleting Snapshots," *CHI 14 Extended Abstracts on Human Factors in Computing Systems* (CHI EA 14), 2014, pp. 2521–2526.
21. T.C. Walber, A. Scherp, and S. Staab, "Smart Photo Selection: Interpret Gaze as Personal Interest," *Proc. SIGCHI Conf. Human Factors in*

Computing Systems (CHI 14), 2014, pp. 2065–2074.
22. I. Mols, E. van den Hoven, and B. Eggen. "Making Memories: A Cultural Probe Study into the Remembering of Everyday Life," *Proc. 8th Nordic Conf. Human-Computer Interaction: Fun, Fast, Foundational* (NordiCHI 14), 2014, pp. 256–265.
23. A. Ceroni et al., "Bridging Temporal Context Gaps Using Time-Aware Re-contextualization," *Proc. 37th Int'l ACM SIGIR Conf. Research & Development in Information Retrieval* (SIGIR 14), 2014, pp. 1127–1130.
24. S. Zerr et al., "Privacy-Aware Image Classification and Search," *Proc. 35th Int'l ACM SIGIR Conf. Research and Development in Information Retrieval* (SIGIR 12), 2012, pp. 35–44.
25. L. Lessig, *Code v2*, electronic edition, 2006; http://codev2.cc.

Wolfgang Nejdl is a professor of computer science and the executive director of the L3S Research Center at the Leibniz University of Hannover, Germany. Contact him at nejdl@l3s.de.

Claudia Niederée is a research group leader in the L3S Research Center at the Leibniz University of Hannover, Germany, and is the coordinator of the European project ForgetIT. Contact her at niederee@l3s.de.

This article originally appeared in IEEE MultiMedia, *vol. 22, no. 1, 2015.*

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.

EIC's Message

Multimedia Search: From Relevance to Usefulness

Alan Hanjalic
Delft University of Technology

There has been an amazing amount of work done and progress achieved in the field of multimedia search over the past two decades. As nicely elaborated on by Lei Zhang and Yong Rui in their recent review of the advances in this field,[1] the developments so far have gone through three main stages: the text-based stage, the content-based stage, and the Web-based stage. In the text-based stage, multimedia search was essentially realized as a trivial extension of the classic, text-oriented information retrieval approach, where text was found in the documents accompanying multimedia items (images, video, and music). In the 1990s, researchers started to explore the possibilities of analyzing the actual (audio-visual) content of multimedia items to automatically infer semantic similarities and textual labels reflecting what is depicted in an image or a video frame and what is audible in a soundtrack. The main motivation for this stage was that manually adding texts to multimedia items was tedious and time consuming. Such texts were also typically one-sided and therefore not always informative enough to help locate the targeted multimedia item.

More recently, inspired by the rapid development of social media platforms, text has come back as an important information source for multimedia indexing and search; however, now it exists in the form of user-generated tags, comments (YouTube), and messages (tweets) accompanying multimedia items being uploaded on such platforms. Although still manually added, this "new text" has become scalable through social interaction. Furthermore, text inserted by various people covers different perspectives, increasing the richness of textual metadata. All this has helped text reenter the field so it can be increasingly exploited by the new wave of approaches marking the Web-based stage, where it is integrated with content-based analysis to make indexing and search more robust and reliable.

Different stages have been marked by the dominance of some key theoretical and algorithmic paradigms. For instance, the content-based stage started with an exploration of a broad range of audio-visual signal analysis methods,[2] but it later increasingly turned more toward well-established algorithmic frameworks, such as support vector machines (SVMs). Similarly, graph theory has provided the main analytic framework for the Web-based stage. The most recent algorithmic hype in the field is deep learning.

If we analyze these three stages and the related (dominating) technologies, we can conclude that their evolution has essentially been resource-driven. In other words, as we have moved from the text-based stage, via the content-based stage, to the Web-based stage, we have included more and more information resources in the development of methods for indexing and searching multimedia. Such technologies have then been investigated and enhanced to get the most out of these resources and improve the results of multimedia search in view of objective criteria, such as average precision (AP).

The key to understanding how a multimedia search system works lies in understanding the notion of relevance. Relevance has so far been explored and optimized mainly with respect to queries. The uncertainty in the channel connecting the query and the collection is typically large, and the need to overcome this uncertainty undoubtedly justifies the tremendous effort invested in the development of various relevance models over the past years,[3] involving more and more information resources and applying new generations of algorithms. An increasing number of results reported in recent literature indicate, however, that pursuing this resource-exploitation target of "technically" optimizing the relevance to a query may have already reached its limits. Specifically, it is unclear how much scientific breakthrough can still be achieved and how much impact all the new proposed solutions will have in a practical technological context.

I believe it is time that we start looking more actively beyond relevance to the query and toward maximizing the usefulness of the search results for the user who inserted the query. This is not as straightforward as it may seem because, to do this, we also need to focus on the uncertainty between the user and the query—that is, to what extent does the query express the information need of the user?

Where does this uncertainty come from? Well, the user always has a reason for inserting a query, namely, a goal to do something with the obtained search results. In other words, there is always a particular search intent underlying the query insertion process. We can type in a query to locate content that might change our mood, for example, to find a funny picture or a video clip that will make us laugh. We may search for a multimedia document that will provide information about a topic of interest or a tutorial on how to do something (such as a video showing how to build a house). Finally, we may also search for a video taken by a head-mounted camera attached to a bungee jumper to help us immerse ourselves in the fantastic experience of a bungee jump, without actually having to jump ourselves.

All these reasons are difficult to express in a typically short query term, and they are also typically not (or insufficiently) encoded in the indexes of multimedia items. This is why query formulations for such examples are often done along the topical dimension to secure a match with a collection. The uncertainty related to the intent remains open, however, and with this also the risk of receiving inadequate search results. Interestingly, because of the topical bias of queries and rather mature solutions to filter out the items from a collection that do not match a topic, the received results are almost always "relevant" because they are related to the query's subject matter. Because of the uncertainty related to the user's intent, however, the returned relevant results are not necessarily useful in view of the user's true information need.

There are several interesting scientific challenges that the multimedia research community needs to face in handling the uncertainty inherent in search intent:

- Why do users search for multimedia items? A comprehensive set of search intent categories is needed.
- How do we assess whether a given multimedia item fits in a given intent category?
- How do we infer the proper intent category from the search session?
- How do we organize the search results to optimally fit the search intent?

The first explorative studies related to these challenges have already been reported in recent literature.[4,5] However, there is still a long road ahead of us to really understand users' needs in multimedia search and to develop systems that can optimally match these needs or, to formulate it more bluntly, to come up with multimedia search solutions that truly matter.[6] I sincerely hope to see more works along this new direction in the future. **MM**

References

1. L. Zhang and Y. Rui, "Image Search—From Thousands to Billions in 20 Years," *ACM Trans. Multimedia Computing, Comm. and Applications (TOMCCAP)*, vol. 9, no. 1, 2013, article no. 36.
2. A. Hanjalic, R.L. Lagendijk, and J. Biemond, "A New Key-Frame Allocation Method for Representing Stored Video Streams," *Proc. 1st Int'l Workshop on Image Databases and Multimedia Search*, 1996, pp. 64–67.
3. L. Yang, "Advanced Retrieval Models for Web Image Search," doctoral thesis, Delft Univ. of Technology, 2013.
4. A. Hanjalic, C. Kofler, and M.A. Larson, "Intent and its Discontents: The User at the Wheel of the Online Video Search Engine," *Proc. 20th ACM Int'l Conf. Multimedia*, 2012, pp. 1239–1248.
5. C. Kofler, M.A. Larson, and A. Hanjalic, "Intent-Aware Video Search Result Optimization," *IEEE Trans. Multimedia*, vol. 16, no. 5, 2014, pp. 1421–1433.
6. A. Hanjalic, "Multimedia Retrieval that Matters," *ACM Trans. Multimedia Computing, Comm. and Applications (TOMCCAP)*, vol. 9, no. 1s, 2013, article no. 44.

Alan Hanjalic is an associate editor in chief of *IEEE MultiMedia* and a professor of computer science and the head of the Multimedia Computing Group at the Delft University of Technology, The Netherlands. Contact him at a.hanjalic@tudelft.nl.

This article originally appeared in IEEE Multimedia, *vol. 22, no. 1, 2015.*

Editors: Gitta Domik and Scott Owen

Preparing Undergraduates for Visual Analytics

Ronald A. Rensink
University of British Columbia

Visual analytics (VA) combines the strengths of human and machine intelligence to enable the discovery of interesting patterns in challenging datasets.[1] Historically, most attention has been given to developing the machine component—for example, machine learning or the human-computer interface. However, it is also essential to develop the abilities of the analysts themselves, especially at the beginning of their careers.

For the past several years, we at the University of British Columbia (UBC)—with the support of The Boeing Company—have experimented with various ways of preparing undergraduate students for VA. Although inspired by the need to prepare students to become visual analysts, the result turned out to be fairly general in scope, applicable to other analytical approaches, as well as more general research. In hindsight, this makes considerable sense. Although the visual component of VA is necessary, it is insufficient; many analytical activities at the human end involve nonvisual skills, such as effective decision-making and the ability to quickly focus on the relevant parts of a problem.

The result of this experimentation is a third-year undergraduate course titled Cognitive Systems 303 (COGS 303) that focuses on "VA unplugged"—that is, on developing investigative abilities prior to training on the VA systems themselves. It was felt that if students focused on developing basic analytical habits of thought prior to learning VA systems, these habits would be reinforced by subsequent practice on "live" systems. As such, this course can serve as preparation for a more technically based VA undergrad or graduate course, such as that of Niklas Elmqvist and David Ebert.[2] Indeed, the course content could be considered an expansion of the initial few weeks spent by Elmqvist and Ebert's course on "analytical reasoning." Combining this with a more technically based course is the main way of teaching VA at UBC, and it is currently being considered as an element of various professional master's degree programs.

COGS 303

COGS 303 is a standard, three-credit, 12-week course open to any upper-level undergraduate in the constituent departments of UBC's interdisciplinary Cognitive Systems Program: computer science, linguistics, philosophy, and psychology. It has been offered—and continually revised—twice each year since the fall of 2008. Classes typically have about 30 students, although larger sizes could be possible. COGS 303 assumes a second-year background in statistics and a second-year course in human perception and cognition, but nothing beyond that. It is offered in the third year to balance two competing factors: students must be sufficiently mature to learn the analytic techniques, but young enough to apply them in their undergraduate career. Application of the analytic skills learned is typically via a subsequent directed-studies course done in a lab or nonprofit organization or, because many of the skills apply to research in general, subsequent courses on research work.

The exact form of the course has varied slightly over the years; Table 1 shows a typical syllabus. The main course textbook is *Asking the Right Questions* by M. Neil Browne and Stuart Keeley.[3] (See the "Additional Reading" sidebar for the other required materials.)

As is evident, the course has three sections, each focusing on a different level of skill. Each section has eight classes, which at a rate of two per week (at 80 minutes a class), require about a month to cover. These sections are crosscut by two recurring themes: the need for continual improvement and the need for effective communication. These themes would likely help in any course, but they are especially relevant in the training of analysts.

The course has been a favorite of most of the students who have taken it, receiving an average rating of 4.4 on a five-point scale. Of the students that eventually became analysts, all those who were asked reported that it played an essential role in their professional development.

Table 1. Typical syllabus for COGS 303.*

Date	Essay	Topic	Readings
0. Basics			
Week 1		Intro to the course, mindsets	Dweck
Section 1: Evaluating an Argument			
Week 2a	1-A	Belief versus knowledge	Burton; B&K, chap. 1
Week 2b	1-B	Structure of arguments	B&K, chap. 4; Booth et al.
Week 3a	–	Meanings of terms	B&K, chap. 5; Niederman & Boyum; Freedman
Week 3b	Debates I		
Week 4a	2-A	Assumptions	B&K, chap. 6; Heuer, chap. 6
Week 4b	2-B	Reasoning	Taleb; B&K, chap. 7
Week 5a	–	Evidence, recap	B&K, chap. 8, pp. 106–116
Week 5b	Debates II		
Section 2: Finding an Explanation			
Week 6a	3-A	Observation and mindset	Heuer, chap. 2; Beveridge, chap. 8
Week 6b	3-B	The role of hypothesis	Heuer, chap. 4; Beveridge, chap. 4
Week 7a	–	Intuition	Beveridge, chap. 6; Claxton
Week 7b	Debates III		
Week 8a	4-A	Imagination, visualization	Beveridge, chap. 5; Brown
Week 8b	4-B	The role of chance	Gilovich; Abelson, pp. 1–11
Week 9a	–	Choosing among alternatives, recap	Cadsby; Heuer, chap. 8
Week 9b	Debates IV		
Week 10a	Midterm exam		
Section 3: Systematization			
Week 10b	–	Experiment design	Cohen
Week 11a	T1-B	Comparative analysis	Yoon, Mayr, Kruskal & Wish
Week 11b	T1-A	Hierarchies, networks	Simon; Barabási, pp. 55–64
Week 12a	T2-B	Power laws, 80/20 rule	Barabási, pp. 65–73, 79–92
Week 12b	T2-A	Research questions, recap	Meltzoff, pp. 13–30; Abelson, pp. 11–14

** Tn = target essay n (an essay along with a presentation about the topic covered). A, B = cohort (the part of the class handing in an essay or presenting that day). B&K = Browne and Keeley textbook.[3] See the "Additional Reading" sidebar for the other required materials.*

Analytic Skills

The heart of COGS 303 is the development of useful analytic and research skills. This is done in three stages. Each is handled via a set of classes (or section) that aims to create a "layer" of particular skills, which is subsequently incorporated into later ones (see Table 1).

Most classes begin with a brief quiz on the readings assigned for that day. (Students are assumed to have done these at home.) The importance of the topic is then briefly discussed so as to provide further motivation. The remainder of the class is devoted largely to group work that exercises the skills covered. Most classes involve the analysis of some essay by each group; many include additional exercises for reinforcement.

Section 1: Evaluating an Argument

The first section focuses on critical thinking. Material is based on the Browne and Keeley textbook,[3] although other texts covering similar material could have been used. Topics include the following:

- *Belief versus knowledge.* The goal here is to shake the conviction of students that their beliefs are necessarily correct. Case studies and neurological reports show that the degree of intensity of a belief does not always correspond to the likelihood that it is true.
- *Structure of arguments.* Students are next shown that the way to higher-quality knowledge is via careful, structured argument. Material is presented concerning the structure of arguments (such as premises and conclusions, role of assumptions, and descriptive versus prescriptive arguments).
- *Meanings of terms.* Here, the goal is to show students that the meanings of words are not always

Additional Reading

The following list of materials details the sources for the required reading other than the main Browne and Keeley textbook.[3] The order is that in which they are encountered by students.

- C.S. Dweck, *Mindset: The New Psychology of Success*, Random House, 2006, pp. 3–11, 173–177.
- R.A. Burton, *On Being Certain*, St. Martin's Press, 2008, pp. 7–20.
- W.C. Booth, G.G. Colomb, and J.M. Williams, *The Craft of Research*, 2nd ed., Univ. of Chicago Press, 2003, pp. 114–123.
- D. Niederman and D. Boyum, *What the Numbers Say*, Broadway Books, 2003, pp. 60–63.
- D.H. Freedman, *Wrong: Why Experts Keep Failing Us—And How to Know When Not to Trust Them*, Little, Brown, & Co., 2010, pp. 37–41.
- N.N. Taleb, *Fooled by Randomness*, Random House, 2004, pp. 185–203.
- G. Claxton, *Hare Brain, Tortoise Mind*, Ecco Press, 1997, pp. 85–95.
- J.R. Brown, *The Laboratory of the Mind*, Routledge, 1991, pp. 1–7
- T. Gilovich, *How We Know What Isn't So: The Fallibility of Human Reason in Everyday Life*, Free Press, 1991, pp. 9–28.
- R.P. Abelson,*Statistics as Principled Argument*, Erlbaum, 1995, pp. 1–11.
- T. Cadsby, *Closing the Mind Gap: Making Smarter Decisions in a Hypercomplex World*, BPS Books, 2014, pp. 88–98, 123.
- P.R. Cohen, *Empirical Methods for Artificial Intelligence*. MIT Press, 1995, pp. 1–10, 67–79.
- C.K. Yoon, *Naming Nature*, Norton, 2009, pp. 5–10.
- E. Mayr, *The Growth of Biological Thought*, Belknap, 1982, pp. 30–32.
- J.B. Kruskal and M. Wish, *Multidimensional Scaling*, Sage, 1978, pp. 7–16.
- H.A. Simon, *The Sciences of the Artificial*, 3rd ed., MIT Press, 1996, pp. 183–197.
- A.-L. Barabási, *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*, Penguin, 2002, pp. 55–73, 79–92.
- J. Meltzoff, *Critical Thinking About Research*, APA Press, 1997, pp. 3–30.
- R.P. Abelson, *Statistics as Principled Argument*, Erlbaum, 1995, pp. 11–14.

as clear-cut as usually believed—there is always some mismatch between word and reality. Emphasis is placed on the dangers of ambiguity and equivocation. Operationalization of terms is also covered.

- *Assumptions*. This introduces the various kinds of assumptions that can be made, such as descriptive assumptions (about facts) and prescriptive assumptions (about values). Students are shown that definitions can also contain assumptions.
- *Reasoning*. Students are shown that their "natural" thinking is not always rational—explicit training is essential. Also covered is the distinction between System 1 (reflexive thinking) and System 2 (effortful thinking) proposed by Daniel Kahneman and Amos Tversky.[4]
- *Evidence*. Students develop a feel for what is good evidence and what is bad, and they learn to avoid standards that vary according to personal beliefs. This includes the effects of personal bias, appeals to authority, the preference for the concrete over the abstract, and various fallacies concerning the believability of statistical studies.

Section 2: Finding an Explanation

Once a basic layer of critical thinking is in place, the next step is to develop skills for evidence-based investigation. Material is largely drawn from two texts. This first is *The Psychology of Intelligence Analysis* by Richards Heuer,[5] which describes the strengths and weaknesses of human cognition when applied to the intelligence domain; the second is *The Art of Scientific Investigation* by W.I.B. Beveridge,[6] which covers the skills used in scientific research, such as observation, intuition, and imagination. (In some ways, the content of COGS 303 could be viewed as a successor to Heuer's book, covering many of the same concerns, but drawing upon more recent developments in human cognition and involving a somewhat broader range of issues.) And in spite of the difference in the domains covered by the two books, there is a strong overlap of the skills discussed, which helps drive home the point that analytical techniques can be general in their applicability.

The topics in this section include the following:

- *Observation and mindset*. Students are made aware that perception is not passive, that what they observe is strongly influenced by what they expect (which in turn is based on what they know). Examples include the distortion of evidence to fit beliefs and the difficulty in changing one's mind so as to see new things.
- *The role of hypothesis*. As a follow up, the goal here is to show that hypotheses can help cope with incomplete information, but they must be used carefully. Discussion includes the strengths and weakness of three different kinds of hypotheses: situational analysis, theory, and analogy.
- *Intuition*. The goal here is to show that the nonconscious mind is a distinctive form of intelligence with particular strengths and weaknesses. (Explicit connections are made to the System 1 and System 2 distinction introduced earlier.) Discussion and exercises include how to effectively

coordinate the use of both kinds of intelligence.

- *Imagination, visualization*. Here students are shown the power of visual thinking, which is portrayed as a distinct form of intelligence with its own strengths and weaknesses. Applications discussed include the use of VA systems to discover new hypotheses. (The visualizations introduced here are also used in later classes when discussing relevant points.)
- *The role of chance*. The goal here is to have students realize that the patterns they detect are sometimes only the result of chance. Therefore, more objective tests (for example, those based on statistics) must be used before asserting conclusions with any confidence.
- *Choosing among alternatives*. Students are taught to make more effective decisions via the Analysis of Competing Hypothesis technique used by intelligence analysts.[5] This is broadened to include the case of several alternatives—not only competing hypotheses, but competing options in general.

Section 3: Systematization

After the basic analytical skills have been developed, attention turns to advanced strategic skills to guide higher-level thinking. This involves a rather heterogeneous set of topics; some focus on developing a feel for complex systems, while others are concerned with more conventional topics such as experiment design. In accord with its heterogeneity, the material for this section is drawn from a variety of sources (see Table 1 and the sidebar). The following topics are covered:

- *Experiment design*. The goal here is to give students a feel for the kind of data to obtain when data either isn't available or is unsuitable (for example, the need to establish causation rather than just correlation). This section also includes discussion of exploratory versus confirmatory analysis.
- *Comparative analysis*. Students are introduced to various ways of investigating complex systems when a controlled experiment is not possible. Examples include natural experimentation (where variation occurs naturally) and classification (where correlations across several dimensions create groups).
- *Hierarchies, networks*. The goal here is to make students aware that some nonlinear systems can have an interesting structure (such as hierarchies) and can sometimes have counterintuitive properties (such as small-world connections).
- *Power laws, 80/20 rule*. Students are introduced to another important characteristic of interacting systems: power-law distributions of properties. This includes a discussion of how to use such distributions (when they exist) to become effective in various activities.
- *Research questions*. The focus here is on finding a question for which the answer will make an important difference. Although of primary interest to researchers, this topic is also somewhat relevant to analysts, if only to help determine which issues to consider.

Crosscutting Themes

In addition to analytic skills, two other kinds of skills are developed: those that enable the student to continually improve their performance (including their intelligence) and those involving communication. These are handled via two crosscutting themes, usually covered via the exercises carried out in each class.

Theme 1: Continual Improvement

This theme is largely drawn from the work of Carol Dweck,[7] who argues that to learn more effectively students must move away from viewing intelligence as a static, fixed quantity (like eye color) and begin to view it as a dynamic quantity that can increase with practice (like strength). Students are encouraged to "play around" with problems and not overly worry about making errors or looking silly—the goal is not to avoid mistakes at all costs, but to learn from them.

The theme of continual improvement is sufficiently important that much of the introductory week is devoted to it. Much of the information that students use for improvement is obtained via peer evaluation in class; such peer-instruction techniques have been shown to help students consolidate new material and apply it to concrete situations.[8]

Theme 2: Effective Communication

The second theme concerns communication skills. These are essential to analysts, if only to enable them to clearly understand the material they encounter and to report their results clearly and concisely. Two forms of communication are emphasized:

Writing. Students are asked to write and then analyze numerous essays, with the topic being the focus of the class that day. A 10-minute session on writing is given at the end of the first few classes, based on the classic grammar book by Strunk & White.[9] Each student is asked to write six essays, each of no more than 400 words, generally at the

This article originally appeared in IEEE Computer Graphics and Applications, *vol. 35, no. 2, 2015.*

rate of one every other week. The essay topic usually involves the material under consideration for that class. The limit of 400 words is severe, but it motivates students to be concise.

Analysis is done by groups of four to five people, and the goal is to have students learn from their peers, allowing best practices to spread. Each group is given three copies of an essay from a classmate and then asked to determine its two biggest weaknesses, in terms of both content and style, and its main strength. Each essay (and group analysis) is marked by an instructor or a teaching assistant. Although only a fraction of the essays submitted can be analyzed by a group during a given class, an effort is made to ensure that each student gets feedback from at least a few groups over the duration of the course.

In accord with the theme of continual improvement, groups are also asked to suggest improvements for the problematic parts. Borrowing a technique commonly used in writers' workshops, the essays have no names attached—instead, only a nonpersonal identifier is used (such as the name of a superhero) known only to the student and the instructors. Thus, groups need not deal with any awkward issues that might arise from analyzing the work of someone they may know. On occasion, a group may receive an essay written by one of its own members. Students are asked not to give away their identities so they can hear what others really think of their work. Students often report this to be a rather educational moment.

Verbal presentation. To help develop this skill (as well as practice their analytic skills), teams of two to three students occasionally debate each other on a selected topic. To encourage conciseness, each student must give their address (opening statement, rebuttal, or summary) in less than three minutes. Participants are given written feedback by the other students regarding content (for example, how well the topic was addressed) and style (for example, how clear the presentations were). To encourage mental flexibility, a team is not assigned its position on the given issue—pro or con—until a coin toss just before the debate begins.

Once students have done a few debates, they are asked to give individual presentations about their target essays. This involves a verbal delivery along with a set of slides. As in the debates, the audience gives written feedback. To encourage students to relax, this feedback is qualitative only (for example, whether the student spoke too quickly) and does not affect their mark, which derives from the essay they wrote. Presentations are again limited to three minutes, in part to allow several students to be assessed in each class and in part to encourage people to be brief. At the end of each presentation, the class is asked to give the presenter a standing ovation, no matter what, to honor the effort made.

The modular design of COGS 303 allows it to be adapted in various ways. For example, if students are already familiar with critical thinking, the first section could easily be dropped and more emphasis could be placed on other topics (such as separate classes for imagination and visualization or more time on logical fallacies). If time is limited, section 2 (along with a few classes from section 3) could be inserted into a more technically based course to provide a nice combination of basic VA and analytic skills.

For more information about this course, visit the COGS 303 website at www.cogsys.ubc.ca/303/.

References

1. J.J. Thomas and K.A. Cook, eds., *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, IEEE CS, 2005.
2. N. Elmqvist and D.S. Ebert, "Leveraging Multidisciplinarity in a Visual Analytics Graduate Course," *IEEE Computer Graphics & Applications*, vol. 32, no. 3, 2012, pp. 84–87.
3. M.N. Browne and S.M. Keeley, *Asking the Right Questions: A Guide to Critical Thinking*, 11th ed., Pearson Education, 2013.
4. D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus and Giroux, 2013.
5. R. Heuer, *The Psychology of Intelligence Analysis*, Center for the Study of Intelligence, 1999.
6. W.I.B. Beveridge, *The Art of Scientific Investigation*, Blackburn Press, 2004. (Originally published 1957.)
7. C.S. Dweck, *Mindset: The New Psychology of Success*, Random House, 2006.
8. A.P. Fagen, C.H. Crouch, and E. Mazur, "Peer Instruction: Results from a Range of Classrooms," *The Physics Teacher*, vol. 40, no. 4, 2002, pp. 206–209.
9. W. Strunk Jr. and E.B. White, *The Elements of Style*, 4th ed., Longman, 1999.

Ronald A. Rensink *is an associate professor in the Departments of Computer Science and Psychology at the University of British Columbia. Contact him at rensink@cs.ubc.ca.*

Contact department editor Gitta Domik at domik@uni-paderborn.de and department editor Scott Owen at sowen@gsu.edu.

Editors: Cláudio T. Silva, csilva@nyu.edu, and Daniel Weiskopf, weiskopf@visus.uni-stuttgart.de

Parallel Coordinates for Multidimensional Data Visualization:

Basic Concepts

Julian Heinrich | Commonwealth Scientific and Industrial Research Organisation
Daniel Weiskopf | University of Stuttgart

Multidimensional data is regularly produced by simulations or measurements in virtually any scientific or engineering field. One example is the set of data attributes computed in computational fluid dynamics (CFD): such simulations typically include the velocity field of the fluid along with temperature, pressure, and other quantities of the flow. Other examples include amino acid attributes of protein structures in molecular biology or multidimensional state variables gathered from sensor networks.

These examples show that it's important to understand the distribution and relationship of data values—even if we ignore their connection to the original 2D or 3D simulation domain.

Standard visualization techniques for data attributes include histograms (for one data dimension) and scatterplots (to show two data dimensions simultaneously). But what if we want to show more than two data dimensions? Then we can split them, showing them side by side. We can look at histograms individually or put several histograms for different data dimensions next to each other. Or we can use a collection of scatterplots for varying pairs of data dimensions. We can even do this systematically, putting all combinations of two data dimensions in a scatterplot matrix.

But this approach has a serious weakness: it doesn't show the connection between more than two data dimensions. Here, parallel coordinates come into play because they aren't restricted to displaying just two data dimensions. In fact, they're scalable in the number of dimensions visualized—we just add another parallel axis for any additional data dimension.

The notion of parallel coordinates dates back to Maurice d'Ocagne in 1885.[1] Parallel coordinates as a means of multidimensional data visualization were developed and popularized by Alfred Inselberg[2] and Edward Wegman.[3] In this article, we describe the basic concepts of parallel coordinates as a tool for multidimensional data visualization. To this end, we discuss the geometric construction of parallel coordinates, issues related to the visualization of large data typical for scientific and engineering disciplines, and how we can interact with the visual representation of multidimensional datasets.

We'll describe details of implementations and applications in a follow-up article in this department. For background reading on parallel coordinates, we recommend Inselberg's book[4] and our state-of-the-art report.[5]

Geometry of Parallel Coordinates

How do we construct a parallel-coordinates plot? Let's start with a simple case: just two data dimensions. Figure 1 shows how a 2D data point is transformed from the domain of the scatterplot to the corresponding line in the domain of the parallel coordinates. The scatterplot domain is spanned by the two orthogonal axes of the two data dimensions, whereas the data axes are parallel in the domain of the parallel coordinates. The data point is mapped to a line in the parallel-coordinates domain so that this line intersects the two parallel-coordinates axes at the respective data values for the two data dimensions. If there's more than just one data point, we just have to draw several lines in parallel coordinates.

Now we know how to map a point from the scatterplot to parallel coordinates. Can we also transform points from parallel coordinates back to the scatterplot? In fact, we can. Figure 1a illustrates how a collection of points coincide on

 Published by the IEEE Computer Society

a line in the scatterplot. When we transform these points to parallel coordinates, they lead to a set of lines that intersect in a single point (Figure 1b). Therefore, we can relate this common intersection point in parallel coordinates to the line in the scatterplot—in other words, this point is transformed back from parallel coordinates to the scatterplot. There's a correspondence between points and lines in both ways: the point-line duality. (For lines with a positive slope in the scatterplot domain, the corresponding pattern in parallel coordinates is another set of lines; see Inselberg's book[4] for details on how to establish the duality for this special case.)

So far, we've restricted parallel coordinates to two data dimensions. However, it's simple to extend them to n dimensions. Figure 2 shows that we just put more and more parallel axes next to each other. Each pair of neighboring parallel axes is handled as we explained above; here, we clip the line that corresponds to a data point so that the line segment stays within the bounds of the two parallel coordinates axes. Next, we construct a straight line segment for each pair of neighboring axes. This leads to a polygonal line that intersects the parallel coordinates axes at the respective data values. In other words, an n-dimensional data point is mapped to a polygonal line in parallel coordinates.

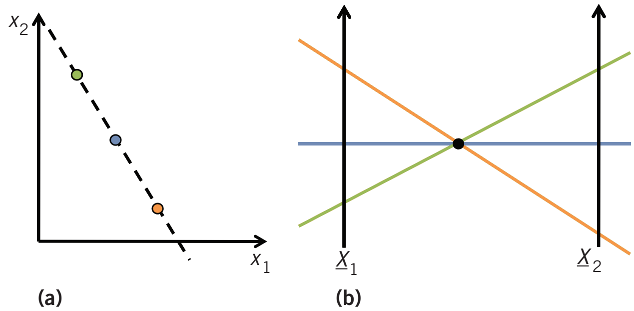


Figure 1. Geometry of parallel coordinates. (a) Data points in a scatterplot correspond to (b) lines in parallel coordinates: these lines intersect the two parallel axes of the parallel coordinates at the respective data values. Point-line duality is established as points coinciding on a line (dashed) in the scatterplot map to lines that intersect at a point (solid black dot) in parallel coordinates.

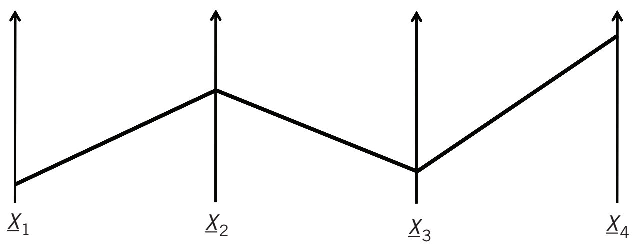


Figure 2. Parallel coordinates for multiple data dimensions and one data point, shown in the form of a polygonal line.

Visual Patterns

We've seen how we can transform back and forth between scatterplots and parallel coordinates. How can we use the geometry of the parallel coordinates to visualize and analyze multidimensional data? How can we read and interpret a parallel-coordinates plot?

Let's have a look at typical visual patterns that you may find in parallel coordinates. First, we restrict ourselves to patterns in 2D parallel coordinates; examples with more data dimensions will follow later. The most striking pattern is the accumulation point in parallel coordinates when we plot data with negative linear correlation. Perfect linear correlation leads to a single accumulation point, following the point-line duality (Figure 3a). However, even more realistic cases with just approximate linear correlation still lead to some smeared-out accumulation region (see Figure 3b). We just need to search for such regions with many line intersections to spot negative correlation.

In contrast, positive correlation leads to parallel or fan-shaped lines in parallel coordinates (Figure 3c). This is a less striking visual pattern and thus can be more difficult to perceive than negative correlation,[6] but we can nevertheless use it to understand the data.

Nonlinear data behavior leads to other visual patterns. These are sometimes easier to spot by observing the envelope of lines, which can be an indicator for further dualities based on the point-line duality, such as curve-curve dualities.[4] A prominent example is the ellipse-hyperbola duality (Figure 3d): we can search for hyperbola patterns in parallel coordinates to find elliptical shapes in the scatterplot that might indicate a spherical or elliptical cluster of data points. Another pattern originates from data consisting of separate clusters: Figure 3e shows an example with two clearly separated clusters, visible in the scatterplot and the parallel-coordinates plot alike.

It's clear that these visual patterns might not be obvious right from the beginning. One issue is that you have to learn how to read parallel coordinates. In fact, with increasing exposure to and experience with working with parallel coordinates, you will become more proficient in identifying relevant visual patterns.

Interacting with Parallel Coordinates

We've seen that negative correlation leads to very strong visual patterns in the form of accumulation points or regions. These patterns are less pronounced for positive correlation.[6] By negating one parallel-coordinates axis, we can turn positive into negative correlation. Therefore, negating is a relevant interaction element. Similarly, the user should be supported in scaling data values to put them into the right perspective.

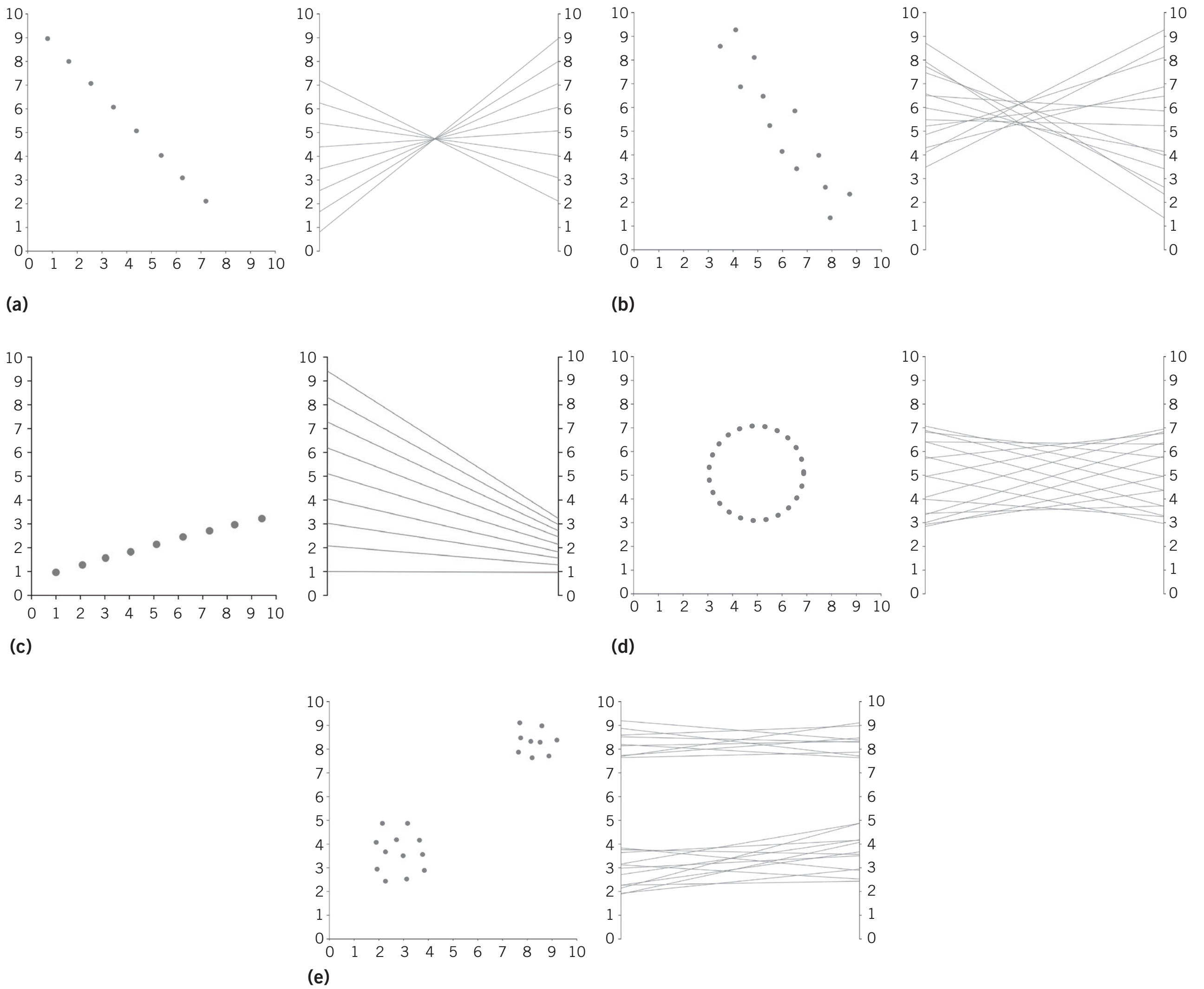


Figure 3. Visual patterns in scatterplots and parallel coordinates: (a) perfect negative correlation, (b) less pronounced negative correlation, (c) perfect positive correlation, (d) ellipse-hyperbola duality, and (e) two clusters.

Let's have a look at *n*-dimensional parallel coordinates. We want to follow the distribution of the data values across data dimensions, essentially reading along the polygonal lines in the plot. It's very useful to select and highlight subsets of the data in parallel coordinates, as in Figure 4, which features data collected by Edgar Anderson.[7] With highlighting, we can easily follow sets of polylines across the data dimensions. The selection is typically done in an interactive fashion; often, the user marks interesting areas in the plot by mouse interaction. This selection process is also called *brushing*. Parallel coordinates come with interesting geometric properties, including the point-line duality. Therefore, we have several options of how we might want to brush data entries. One approach lets us brush data values or ranges thereof on an axis in parallel coordinates: axis-aligned brushing (Figure 5a). An alternative method is angular brushing[8]: a range of angles is selected in parallel coordinates (Figure 5c). In this way, for example, we can highlight positive slopes (correlations) in the data. Yet another approach supports polygon-shaped brushes in the parallel-coordinates domain, corresponding to rotated boxes in the scatterplot domain (Figure 5b).

The full power of brushing emerges in combination with other visualization views, for example, 3D renderings of simulation results. The sidebar "Coordinated Multiple Views" provides more background information.

Even with brushing, it's clear that the visual patterns strongly depend on the order of the coordinate axes.

Correlations between two data dimensions are best shown with a pair of neighboring parallel-coordinates axes. What's the best order of axes? Some approaches use automatic data analysis to identify the most relevant pairs of axes and put them next to each other.[9] Other approaches allow the viewer to change the axis order interactively or provide a set of individual plots to explore all possible pairwise correlations.[10] Even with some (initial) automatic order, any useful parallel-coordinates visualization tool will support the user in modifying the axis order, to interactively explore the data.

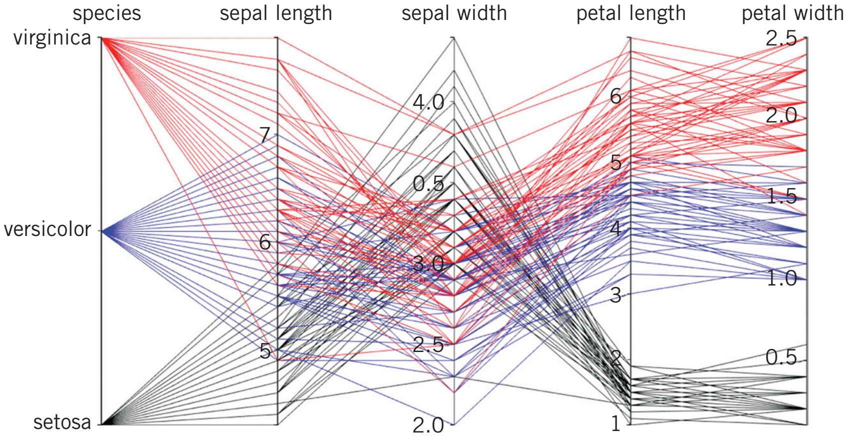


Figure 4. Brushing lets us compare the values of data samples or clusters over a set of dimensions. For this dataset (a popular test dataset of iris flower data collected by Edgar Anderson[7]), a non-metric variable was mapped to a number to represent it in parallel coordinates (leftmost axis). This is common practice, as it allows us to select groups of data samples. We highlighted three different groups with separate colors.

Density Plots

We've been assuming that the actual visual representation of the parallel-coordinates plot is simple: just draw a polygonal line for each data point. More specifically, let's just draw opaque lines. We produced the examples shown so far by using this rendering approach. However, this way of rendering doesn't work for large datasets—large in the sense of many data points. If it is, then the parallel-coordinates plot will suffer from much overplotting and visual clutter. Unfortunately, the issue of large data is natural to most applications in science and engineering because simulations and measurements tend to deliver large numbers of data points.

One popular approach to large-data parallel coordinates replaces the opaque lines with semitransparent ones. In this way, we can see a varying density of lines: high density in regions of many overlapping lines, low density in areas covered by few lines. Such a density plot allows us to easily recognize the prominent visual patterns that belong to the most frequent characteristics of the data, as in Figure 6.

Density plots can be constructed by rendering lines with additive blending: we virtually count the number of times a pixel is covered by a polygonal line from the parallel-coordinates plots. Additive blending is sometimes replaced by alpha blending; this blending model isn't order-independent, that is, it allows us to emphasize the lines that are rendered late. This only occurs if lines have varying density or color. Another alternative uses the model of continuous parallel coordinates[11]: it directly transforms densities from the (spatial) computational domain—such as the simulation grid—to densities in parallel coordinates.

We think that density plots are highly relevant and useful for the typical dataset sizes in science and engineering.

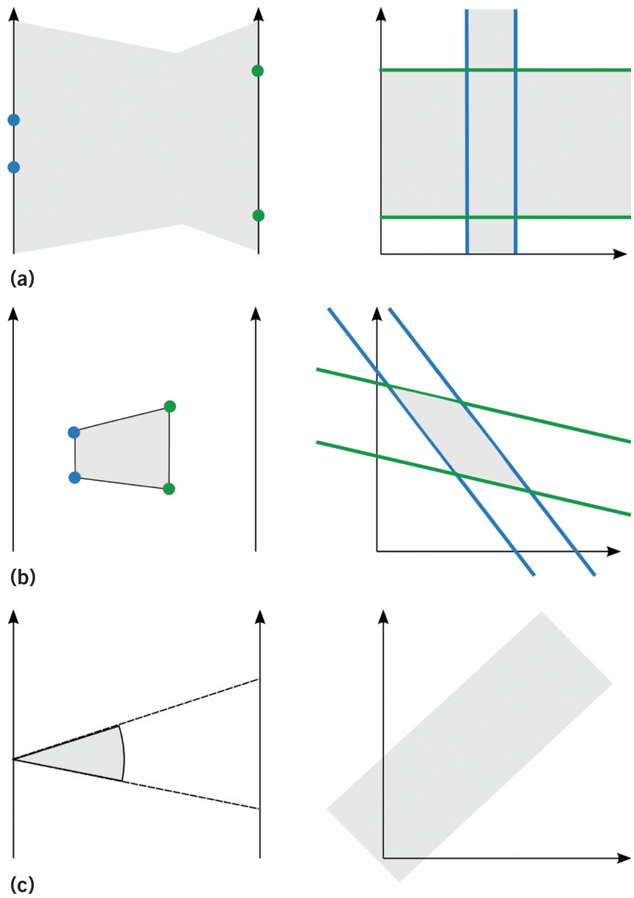


Figure 5. Different types of brushing in parallel coordinates (shown left) and the corresponding selection in the scatterplot (right): (a) axis-aligned brushing, (b) polygon-shaped brush, and (c) angular brushing.

Coordinated Multiple Views

Coordinated multiple views are a popular approach to designing visual interfaces that facilitate the visual analysis of complex data. The basic idea is that complex data can't be shown in a single visualization alone because there's just too much information to be conveyed in a single image. Instead, we can show different types of visualizations in multiple views. The multiple views can provide different perspectives on the very same data. But how can we build a combined, overall mental model of the data from several views? One strategy uses the concept of coordination: the views aren't completely independent from each other, but rather react to user interaction in a coordinated sense.

Brushing-and-linking is a popular approach to coordinated multiple views: visualization elements are brushed (that is, selected) in one view, and the selection is automatically applied to the other views, highlighting the corresponding elements in these views. Early work used brushing-and-linking to coordinate multiple scatterplots.[1] Brushing is commonly used in scatterplots; it can be extended to more complex types of brushes, for example, to high-dimensional brushes.[2]

It's even more interesting to link structurally different views: data in science and engineering often has some intrinsic spatial relationship, and therefore it's especially useful to link different views that complement the perspective of the spatial embedding with the distribution of data values. Figure A illustrates this scenario.[3] Another similar example is brushing-and-linking between scatterplots of data attributes from a simulation and the 3D visualization of the simulation. Helmut Doleisch and his colleagues[4] provide several examples for such coupling. For further reading on coordinated multiple views, we recommend the survey by Jonathan Roberts.[5]

References

1. R.A. Becker and W.S. Cleveland, "Brushing Scatterplots," *Technometrics*, vol. 29, no. 2, 1987, pp. 127–142.
2. A.R. Martin and M.O. Ward, "High Dimensional Brushing for Interactive Exploration of Multivariate Data," *Proc. IEEE Visualization*, 1995, pp. 271–278.
3. J. Heinrich et al., "Visualising Intrinsic Disorder and Conformational Variation in Protein Ensembles," *Faraday Discussions*, vol. 169, no. 1, 2014, pp. 179–193.
4. H. Doleisch, M. Gasser, and H. Hauser, "Interactive Feature Specification for Focus+Context Visualization of Complex Simulation Data," *Proc. Joint Eurographics/IEEE TCVG Symp. Visualization*, 2003, pp. 239–248.
5. J.C. Roberts, "State of the Art: Coordinated & Multiple Views in Exploratory Visualization," *Proc. Int'l Conf. Coordinated & Multiple Views in Exploratory Visualization*, 2007, pp. 61–71.

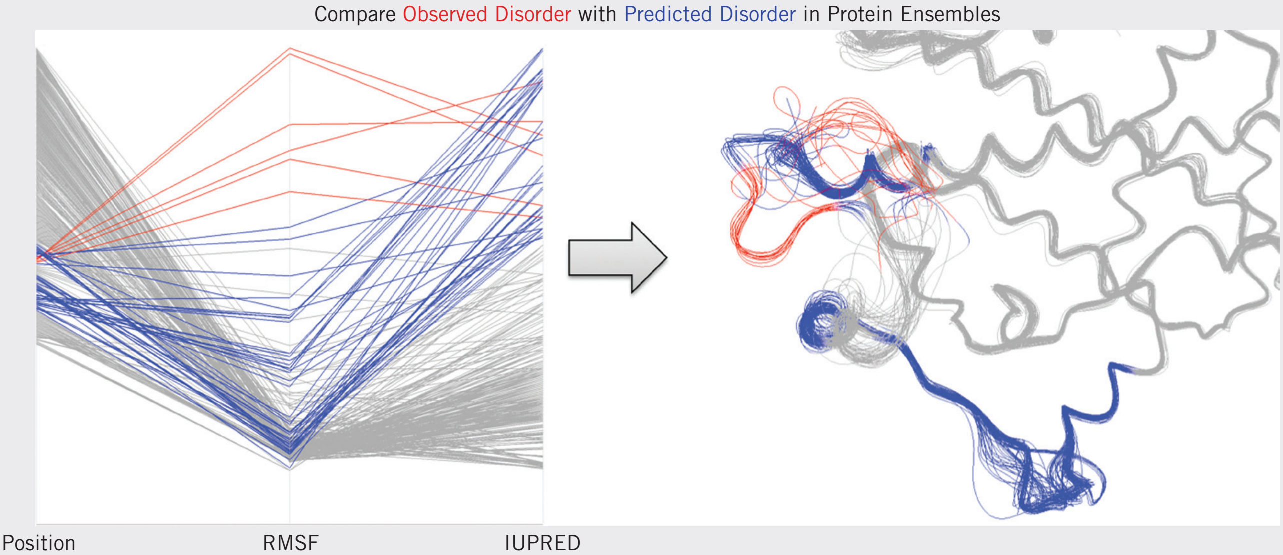


Figure A. Brushing-and-linking using parallel coordinates to highlight amino acids with properties as selected in parallel coordinates (left). The red brush indicates residues of protein structures that show a significant variation in their spatial position, corresponding to high values of the root mean-square fluctuation (RMSF). These residues are called "disordered." Only with linking to the 3D visualization (right) can we understand the spatial relationship of the disordered parts. From the parallel-coordinates plot, we can see that not all residues that are predicted to be disordered (blue brush, IUPRED score > 0.5) show a strong spatial variation (red brush). More details appear elsewhere, in the source for the image.[3]

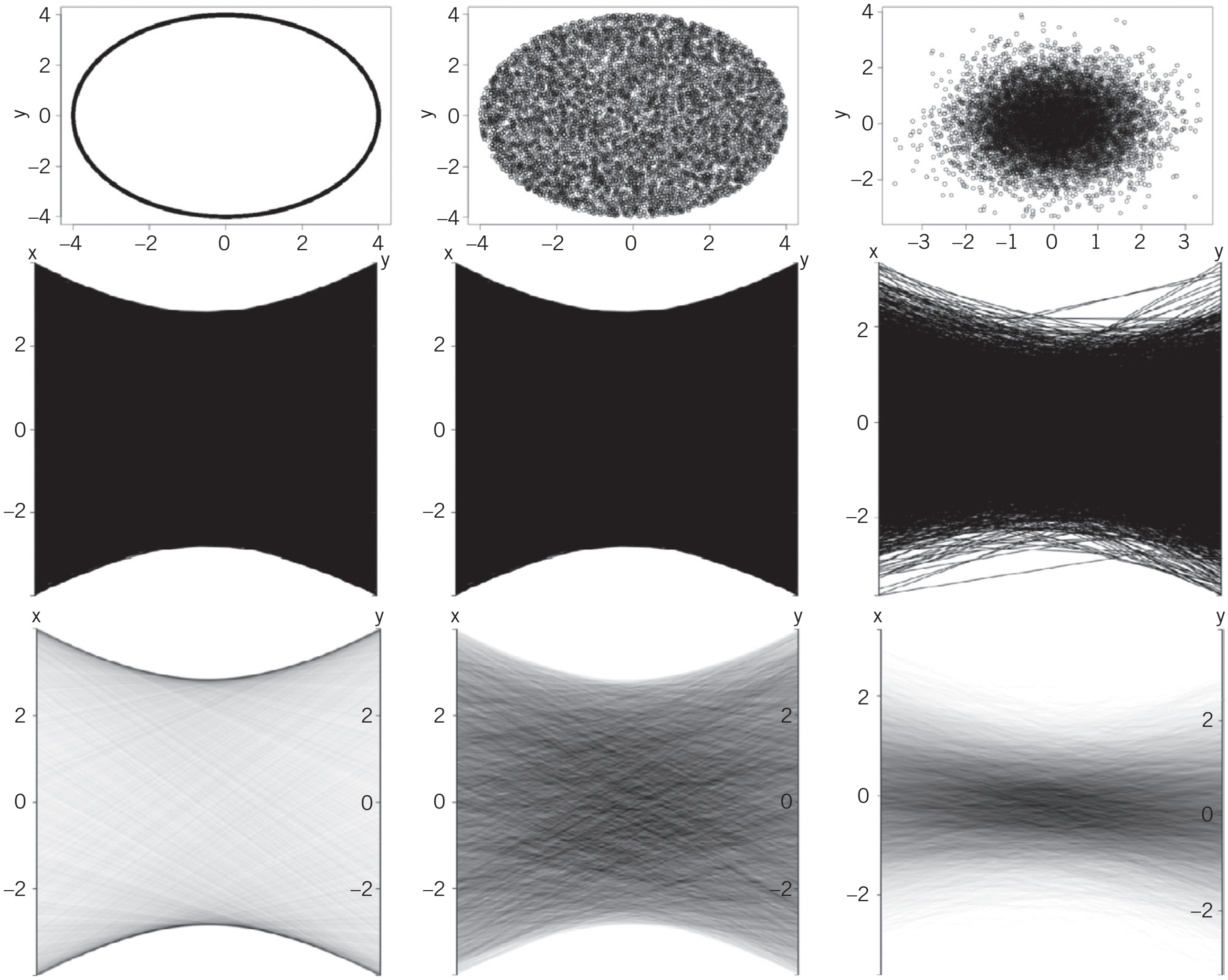


Figure 6. Density plots reveal otherwise hidden patterns in parallel coordinates. The top row shows scatterplots, the center row parallel coordinates with opaque lines, and the bottom row density plots of parallel coordinates. The patterns that emerge in parallel coordinates from samples taken from the circumference of a circle (left column) and from the circle area (center column) are identical for opaque line rendering (center row) but become distinguishable with density information (bottom row). The density plot also provides much more information of the distribution of lines than the opaque line rendering for data distributed with varying density (right column).

However, we would like to point out alternative strategies, such as reduction methods that work on the original data before it's visualized (for example, clustering, principal component analysis, and filtering). More details are in our report.[5]

Parallel coordinates are a great tool for visualizing multidimensional data. With the straightforward mapping from data points (in scatterplots) to lines (in parallel coordinates), the concept and basic implementation of parallel coordinates are simple. So, why aren't plots with parallel coordinates as common as scatterplots? We think that one important reason is that the visual patterns are very different from those in scatterplots and require training before they become useful. A related issue is that some visual patterns are quite hard to perceive—especially the visual asymmetry between visualizing positive and negative correlation is striking and makes some aspects of the data less clear than with scatterplots. We also think that the need for interaction—in particular, for complex data—might add some problems with the acceptance. However, we're convinced that parallel coordinates provide a very useful, complementary view on multidimensional data. This is particularly true for data that comes with relationships between more than two data dimensions.

You can find more background information and material for further reading on Inselberg's webpage (www.math.tau.ac.il/~aiisreal) and www.parallelcoordinates.de. We'll publish a follow-up article in this department with more details of implementations and practical applications of parallel coordinates. Stay tuned! ■

Acknowledgments

The dataset used in Figure 4 shows data collected by Anderson;[7] we used the data distribution made available through the R Project (www.r-project.org).

References

1. M. d'Ocagne, *Coordonnées parallèles et axiales: Méthode de transformation géométrique et procédé nouveau de calcul graphique déduits de la considération des coordonnées parallèlles*, Gauthier-Villars, 1885.
2. A. Inselberg, "The Plane with Parallel Coordinates," *Visual Computer*, vol. 1, no. 4, 1985, pp. 69–91.
3. E.J. Wegman, "Hyperdimensional Data Analysis Using Parallel Coordinates," *J. Am. Statistical Assoc.*, vol. 85, no. 411, 1990, pp. 664–675.
4. A. Inselberg, *Parallel Coordinates: Visual Multidimensional Geometry and Its Applications*, Springer, 2009.
5. J. Heinrich and D. Weiskopf, "State of the Art of Parallel Coordinates," *Eurographics 2013, State of the Art Reports*, 2013, pp. 95–116.
6. J. Li, J.-B. Martens, and J.J. van Wijk, "Judging Correlation from Scatterplots and Parallel Coordinate Plots," *Information Visualization*, vol. 9, no. 1, 2010, pp. 13–30.
7. E. Anderson, "The Irises of the Gaspe Peninsula," *Bulletin Am. Iris Soc.*, vol. 59, 1935, pp. 2–5.
8. H. Hauser, F. Ledermann, and H. Doleisch, "Angular Brushing of Extended Parallel Coordinates," *Proc. IEEE Symp. Information Visualization*, 2002, pp. 127–130.
9. A. Tatu et al., "Automated Analytical Methods to Support Visual Exploration of High-Dimensional Data," *IEEE Trans. Visualization and Computer Graphics*, vol. 17, no. 5, 2011, pp. 584–597.
10. J. Heinrich, J. Stasko, and D. Weiskopf, "The Parallel Coordinates Matrix," *EuroVis–Short Papers*, 2012, pp. 37–41.
11. J. Heinrich and D. Weiskopf, "Continuous Parallel Coordinates," *IEEE Trans. Visualization and Computer Graphics*, vol. 15, no. 6, 2009, pp. 1531–1538.

Julian Heinrich is a postdoctoral fellow at Digital Productivity, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia. His research interests include visualization, visual analytics, and bioinformatics. Heinrich has a PhD in computer science from the University of Stuttgart, Germany. Contact him at julian.heinrich@csiro.au.

Daniel Weiskopf is a professor at the University of Stuttgart, Germany. His research interests include visualization, visual analytics, GPU methods, computer graphics, and special and general relativity. Weiskopf has a PhD in physics from the University of Tübingen, Germany. He's a member of the IEEE Computer Society, ACM, Eurographics, and the Gesellschaft für Informatik. Contact him at weiskopf@visus.uni-stuttgart.de.

This article originally appeared in Computing in Science & Engineering, *vol. 17, no. 3, 2015.*

cn *Selected articles and columns from IEEE Computer Society publications are also available for free at http://ComputingNow.computer.org.*

Editor: **Cesare Pautasso**
University of Lugano
c.pautasso@ieee.org

Editor: **Olaf Zimmerman**
University of Applied Sciences
of Eastern Switzerland, Rapperswil
ozimmerm@hsr.ch

Lightweight and Flexible

Emerging Trends in Software Architecture from the SATURN Conferences

Michael Keeling

Software architecture is a rather young, fast-moving discipline. It's beneficial to reflect once a year about what has been achieved so far, and the SATURN conference has become one of the premier places to do so. Michael Keeling did an excellent job as the SATURN 2014 program chair; in this article he distills four of the most important software architecture trends that were highlighted during the conference: architecting for DevOps, flexible designs, lightweight architecture design methods, and renewed interest in software architecture fundamentals.
—Cesare Pautasso and Olaf Zimmermann, department editors.

OVER ITS 10-YEAR HISTORY, the annual Software Engineering Institute (SEI) Architecture Technology User Network (SATURN) conference has become a barometer for the ever-evolving software architecture climate. SATURN is a great conference, and it was my privilege to be the SATURN 2014 technical chair. Here, I summarize software architecture trends I saw emerge during the conference and give a glimpse of the future based on the current SATURN 2015 technical program. (For a brief rundown of SATURN 2014, see the sidebar.)

Four Trends

If one idea stood out at SATURN 2014, it was that promoting business agility requires sound architecture design. Discussion in the hallways between sessions wasn't about whether architecture and agile worked together but more about how best to achieve agility through architecture. With this as the main theme, I noticed the following four supporting ideas throughout the conference.

Architecting for DevOps

With nine sessions covering DevOps-related topics, DevOps secemed everywhere. DevOps is an emergent software development approach that blends the traditional roles of operations and software development to increase an organization's ability to deliver business value. Groups practicing DevOps take direct responsibility for the whole user experience. This includes everything from infrastructure and tools to development, testing, and deployment, and much else besides.

SATURN 2014 BY THE NUMBERS

Hosted by the Software Engineering Institute (SEI) in collaboration with *IEEE Software*, the 2014 SATURN (SEI Architecture Technology User Network) conference ran from 5 to 9 May in Portland, Oregon. Here's a brief summary of what took place:

- 198 software architects, software developers, and thought leaders attended.
- 27 countries were represented.
- 33 experience report presentations shared stories and lessons across three broad themes: Technical, Methods and Tools, and Leadership and Business.
- 12 tutorials and courses covered topics such as big data, software architecture design analysis, architecture hoisting, DevOps, architecture katas, and architecture-centric agile practices.
- 4 participatory sessions gave attendees the chance to immediately practice new methods.
- 4 keynotes were delivered by Bill Opdyke (JPMorgan Chase), Joe Justice (Scrum Inc. and Team WIKISPEED), Diane Marsh (Netflix), and Jerome Pesenti (IBM).
- 1 panel explored the role of technical debt from the perspective of architecture. The moderator was George Fairbanks (Google); the panel comprised Jeromy Carrière (Google), Philippe Kruchten (University of British Columbia), Robert L. Nord (SEI), Michael Keeling (IBM), and Rebecca Wirfs-Brock (Wirfs-Brock Associates).
- 1 Open Space event was facilitated by Diana Larsen (FutureWorks Consulting), with numerous attendee-led sessions and serendipitous encounters among "butterflies."

Nearly all SATURN 2014 keynotes and experience reports are available online as slides, video, or both.[1,2]

References

1. *SATURN 2014 Presentations*, Software Eng. Inst., 2014; http://resources.sei.cmu.edu/library/asset-view.cfm?assetID=89532.
2. *SATURN: A Software Architecture Community*, Software Eng. Inst., 2014; http://bit.ly/saturn2014-video.

Many of the benefits promised by DevOps can be achieved only when the system's architecture promotes DevOps values and enables specific practices. So, DevOps values must be promoted in the architecture similarly to any desirable quality attribute. Much like security, availability, scalability, and other quality attributes, DevOps can't simply be bolted on as an afterthought.

The importance of architecting for DevOps was evident throughout the conference, the most obvious being during Diane Marsh's keynote describing how Netflix supports continuous delivery. One way Netflix achieves this is through an automated test suite affectionately called the Simian Army. Tools such as Chaos Monkey, a program that randomly kills running application instances in production, codify desired quality attributes in a way that diagrams and prose can't. For example, Netflix's arsenal of automated tools provides practical tests for evaluating how well the system achieves certain runtime quality attributes such as availability and scalability.

In many ways, Chaos Monkey and its companions provide a test-driven-like approach for some quality attributes in the architecture. Chaos Monkey clearly defines reliability and availability in the form of an executable test that continuously runs during normal business hours. A developer's ability to progress through the red–green–refactor cycle at the architectural level is powerful indeed.

Naturally, Netflix didn't arrive at this design overnight. The highly malleable and resilient Netflix architecture is the result of years of evolution and hard-earned lessons. If nothing else, the Netflix design serves as a shining example of what might be achieved when the organization, culture, and technology are aligned.

Flexible Designs

Netflix's experience is also a good example of the importance of flexible designs, and there were many other examples at SATURN. Flexible designs enable organizational agility. Here I discuss two of the more exciting examples from the conference.

Flexibility was essential to IBM's success in the 2011 *Jeopardy!* IBM Challenge. During his keynote, Jerome Pesenti shared the story of IBM's journey to create Watson, a cognitive computing system that defeated the best human *Jeopardy!* contestants. *Jeopardy!* is an American television game show that provides trivia clues as statements, to which contestants provide answers in the form of a question. Watson's victory was an important milestone in AI. However, Watson might not have been possible without the highly flexible, modular, and parallelizable architecture the research team adopted.

Early on, the Watson team realized that no single algorithm would provide question-answering accuracy sufficient to beat the best *Jeopardy!* contestants. The solution lay in integrating multiple algorithms to create compounding improvements in answer accuracy. A highly modular architecture that promoted rapid experimentation was critical to the team's success. By the time Watson was ready to compete, the relatively small 20-person team had conducted more than 5,500 independent experiments over three years, totaling more than 11 million CPU hours.[1] On the basis of these experiments, the team picked the best combination of algorithms for the competition.

As Joe Justice's keynote discussed, flexibility was also essential to Team WIKISPEED's success in the Progressive Insurance X Prize competition. In less than three months, a team of volunteers designed and constructed a car that could achieve over 100 miles per gallon. Justice attributes his team's success to its use of agile methods and object-oriented design. Neither practice is characteristic of automotive manufacturing.

The WIKISPEED garage is built around common agile practices such as pairing, test-driven development, the definition of "done," and Scrum. This lets new teammates quickly pick up new knowledge, apply it, and contribute to vehicle construction. Justice calls the method Extreme Manufacturing, in homage to the method's inspiration, Extreme Programming.

Likewise, WIKISPEED cars are designed to be modular from the beginning. As with the Watson team, modular design lets Team WIKISPEED quickly iterate designs. In fact, WIKISPEED has developed a product line of high-efficiency vehicles, from trucks to sports cars, based on the same modular framework. It's amazing how the fit and function of agile and architecture become plainly evident when the same practices are applied in a setting other than software.

Although Watson and WIKISPEED might be extreme examples of how agile design can achieve extraordinary outcomes, the increasing business demands and complexity of many systems requires that architects think about agility first, regardless of the development methodology.

> Some say that a software architect's first job is to teach.

Lightweight Architecture Design Methods

Trimming the fat from software architecture methods has been a trending theme at the past few SATURN conferences. Topics such as architecturally evident coding styles, lightweight representations, essential modeling skills, and lean design methods have appeared regularly in the SATURN technical program. Many presentations covering these topics have won the conference's *IEEE Software* SATURN Architecture in Practice Presentation Award, including at SATURN 2014. This peer-voted award goes to the presentation that best describes lessons learned while applying architecture-centric practices.

The 2014 award went to Will Chaparro's and my talk, "Facilitating the Mini-Quality Attributes Workshop."[2] I think our presentation won because it provided practical, actionable advice and because the mini-QAW updates the traditional Quality Attribute Workshop[3] to address modern agile teams' concerns.

At the heart of the mini-QAW is a quality attribute taxonomy, a set of common quality attributes relevant to typical stakeholder concerns, classified for easier consumption. Our experience was based on using a taxonomy for enterprise search applications; some generic taxonomies are available in the technical report *Quality Attributes*.[4] By using a taxonomy to enrich facilitation options, we could skip or reduce several steps in the traditional QAW so that the time

FIGURE 1. A quality attribute taxonomy used as a facilitation aid during a mini-QAW (Quality Attribute Workshop) with stakeholders. Sticky notes contain raw scenarios; dots indicate priority determined by a dot-voting exercise.

to complete a workshop decreases from a few days to a few hours, depending on the circumstances.

Mini-QAW facilitators can use a taxonomy directly in many ways. For working with experienced stakeholders, the taxonomy aids structured brainstorming—for example, placing sticky notes on a whiteboard clustered around taxonomic terms (see Figure 1). During a workshop, the facilitator can also "walk the web" with stakeholders by asking questions from a taxonomy-based questionnaire. In this way, the taxonomy provides a ready starting point and a concrete guide for facilitating the workshop. Such a questionnaire is an excellent accompaniment for inexperienced facilitators and stakeholders.

I think lightweight methods have proved popular at SATURN because they enable teams to learn fast, fail fast, change fast, and communicate effectively. These factors are essential for self-organizing teams, which in turn enable better designs to emerge.

Renewed Interest in Software Architecture Fundamentals

None of the many forward-looking trends at SATURN 2014 would have been possible without a strong foundation in software architecture. Teaching the next generation of software architects seemed to be on many speakers' minds. Some of my favorite talks involved either sharing deep knowledge in software architecture or novel approaches to teaching as a means of growing the next generation of architects.

One highlight was Ted Neward's tutorial "Architectural Katas." During an architectural kata, attendees work in small groups to architect a system from scratch, with the facilitator acting as the customer stakeholder.[5] Inspired by code katas (which were in turn inspired by martial-arts katas),[6] an architectural kata creates a no-risk opportunity for software developers and architects, both experienced and novice, to practice software architecture design.

During our kata, teams explored business and architectural drivers and designed several views of a system that showed how the most important quality attributes might be achieved. As a conference organizer, I found it particularly rewarding to see participants applying ideas they had learned at the conference.

New research addressing longstanding ignorance was also great to see. "Approaching Security from an 'Architecture First' Perspective," by Rick Kazman, Jungwoo Ryoo, and Humberto Cervantes,[7] won the *IEEE Software* SATURN New Directions Presentation Award. This peer-selected award goes to the presentation best describing innovative ideas that advance the state of architecture-centric practice. I think this presentation won not because its revelations were necessarily groundbreaking but because Kazman and his colleagues directly addressed longstanding software architecture myths about promoting security. Their research was solid and presented in a way that was relevant to practitioners.

Their most important finding showed that the most effective way to promote security in a system is early adoption of a security framework such as Apache Shiro or Spring Security. In short, "delegating security issues to frameworks allows developers to devote their energy to application logic, increasing overall productivity."[7]

Effectively sharing what we know about software architecture was the focus of George Fairbanks' talk, "Teaching Architecture Metamodel-First."[8] Fairbanks outlined five common obstacles to teaching software architecture, and strategies for overcoming them. For example, to remove abstraction as an obstacle, lessons with new architects should focus first on diagramming rather than analysis and employ plenty of hands-on exercises. Also, to overcome burgeoning architects' tendency to focus on the wrong details, show them how to start with a metamodel by creating the legend for their diagrams first.

Some say that a software architect's first job is to teach. Effectively communicating decisions about a software system's architecture requires that all stakeholders have a firm grasp on software architecture fundamentals. Knowing the fundamentals is important, and so is knowing how to teach them to others. It was great to see talks focused on teaching the teacher, a role every software architect inevitably fills at least sometimes.

SATURN 2015 Highlights

The 11th SATURN conference will be in Baltimore from 27 to 30 April 2015, with the return of both George Fairbanks (the SATURN 2012 technical chair) and me as the technical cochairs.

What trends will emerge from SATURN 2015? Although I don't want to speculate too wildly on what trends might emerge, some talks have already been accepted and offer a glimpse of what might come. Scheduled topics include reducing technical debt, cloud architectures, and microservices. Design thinking and similar reflective, adaptive design frameworks also appear to be trending in the submitted proposals. DevOps is still present but doesn't seem as prevalent as in 2014, but we'll know more as the conference unfurls in April.

The SATURN 2015 keynotes will be by Mary Shaw, a professor of computer science at Carnegie Mellon University and winner of the US National Medal of Technology and Innovation; Mark Schwartz, Chief Information Officer for US Citizenship and Immigration Services; and Gregor Hohpe, Chief IT Architect at Allianz.

Finally, a special track consisting of sessions curated by invited thought leaders will cover a variety of emerging and classic topics. You'll likely recognize many of the speakers' names, including Simon Brown, Ariadna Font, Len Bass, Rebecca Wirfs-Brock, Jeff Patton, Joseph Yoder, and Jeromy Carrière.

For more information and archived presentations, visit www.sei.cmu.edu/saturn/2015. On behalf of the program committee, I look forward to seeing you at SATURN 2015!

References

1. D. Ferrucci et al., "Building Watson: An Overview of the DeepQA Project," *AI Magazine*, vol. 31, no. 3, 2010, pp. 59–79.
2. W. Chaparro and M. Keeling, "Facilitating the Mini-Quality Attributes Workshop," presentation at the 10th SEI Architecture Technology User Network (SATURN) Conf., Software Eng. Inst., 2014; http://resources.sei.cmu.edu/library/asset-view.cfm?assetID=89553.
3. M.R. Barbacci et al., *Quality Attribute Workshops (QAWs), Third Edition*, tech. report CMU/SEI-2003-TR-016, Software Eng. Inst., 2003.
4. M. Barbacci et al., *Quality Attributes*, tech. report CMU/SEI-95-TR-021, Software Eng. Inst., 1995.
5. T. Neward, *Architectural Katas*, Neward & Associates, 2012; https://archkatas.herokuapp.com.
6. D. Thomas, *Code Kata*, blog, 2013; http://codekata.com.
7. R. Kazman, J. Ryoo, and H. Cervantes, "Approaching Security from an 'Architecture First' Perspective," presentation at the 10th SEI Architecture Technology User Network (SATURN) Conf., Software Eng. Inst., 2014; http://resources.sei.cmu.edu/library/asset-view.cfm?assetID=89604.
8. G. Fairbanks, "Teaching Architecture Metamodel-First," presentation at the 10th SEI Architecture Technology User Network (SATURN) Conf., Software Eng. Inst., 2014; http://resources.sei.cmu.edu/library/asset-view.cfm?assetID=89518.

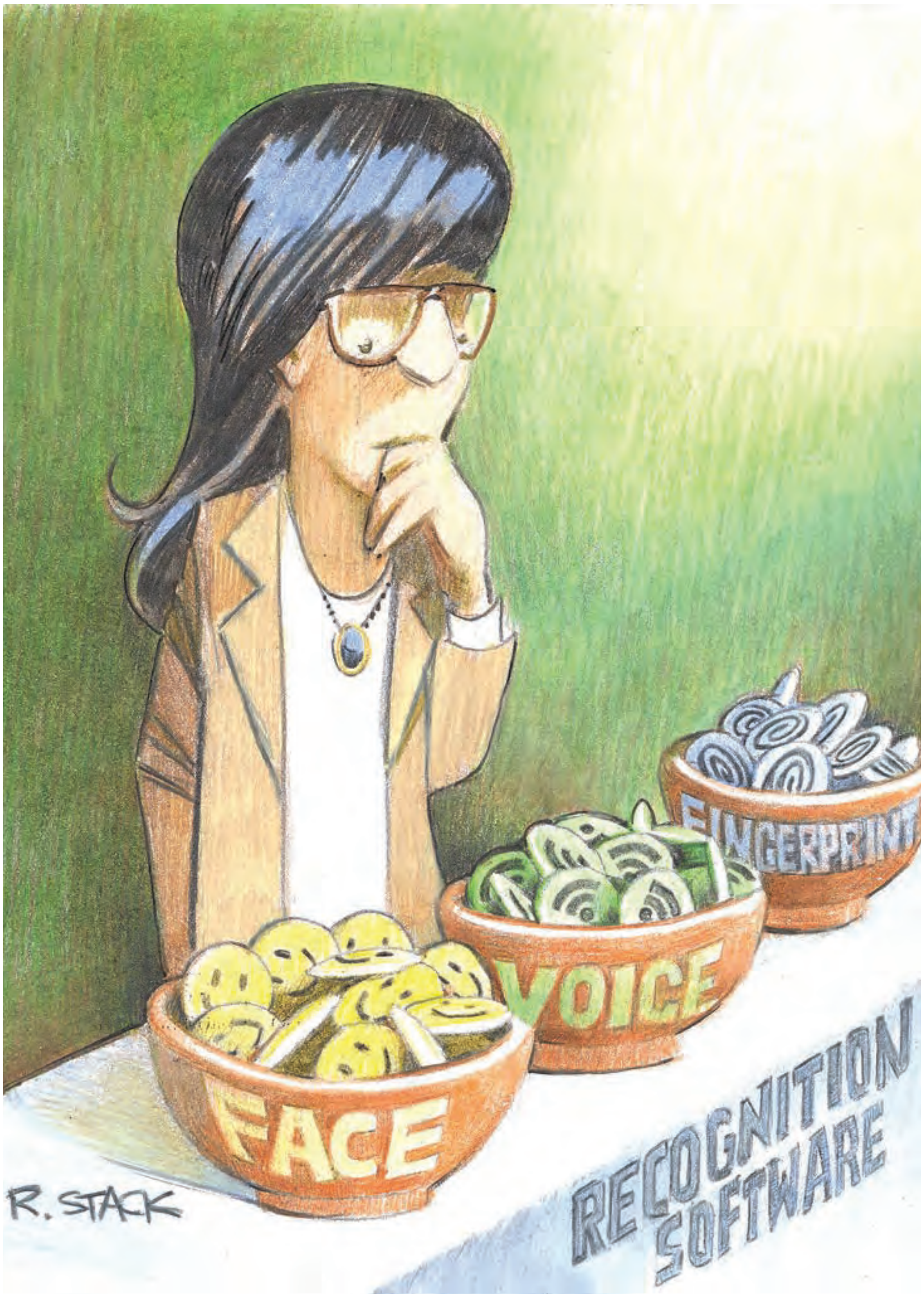
MICHAEL KEELING is a software engineer at IBM and a member of the IBM Watson Explorer team. His professional interests include software architecture design, agile methods, and making awesome software. Keeling received a master's in software engineering from Carnegie Mellon University. Contact him at mkeeling@neverletdown.net.

This article originally appeared in IEEE Software, vol. 32, no. 3, 2015.

Editors: Richard Ford, rford@se.fit.edu | Deborah A. Frincke, debfrincke@gmail.com

Biometric Authentication on Mobile Devices

Liam M. Mayron | Arizona State University



Mobile devices provide a compelling platform for computing. They are small, portable, and increasingly powerful, and many have multicore processors and high-resolution displays. More and more users are performing tasks on smartphones that they might have previously executed on personal computers. These tasks range from composing and sending emails to playing games and everything in between. Smartphones have become the gatekeepers to our most personal information, including our heart rate and bank account information. It's critical to protect the sensitive data available through these devices. One way to do this is through the use of biometrics.

Smartphones contain many useful sensors, such as GPS, cameras, touchscreens, gyroscopes, and microphones; some even have fingerprint sensors. Such sensors allow the use of biometrics as an authentication mechanism on mobile devices.

Authentication on Mobile Devices

Users often decline to use passwords or other methods of securing their mobile devices.[1] Passwords have historically been used to authenticate users; they're convenient to implement, requiring minimal hardware. Passwords require an exact match, which allows them to be protected using one-way transforms, or hashes. They can be authenticated very quickly and provide no ambiguity—a password is either an exact match and is authenticated, or it isn't an exact match and is denied.

Strong passwords are difficult to guess. Generally, longer passwords that don't contain dictionary words but instead are made up of numbers, special characters, and a combination of uppercase and lowercase letters will be more secure. On desktop systems, complex passwords can be difficult for users to remember but are only slightly less convenient to input than simple passwords. Mobile devices, however, have challenged the usability of text passwords. Many smartphones rely on a finger-operated touchscreen as their main input method. Touchscreens are versatile, but they impede the precise input that a physical keyboard provides. Algorithms that effectively autocorrect text can't be applied to passwords, slowing down the authentication process.

Mobile devices are intended for quick, frequent access, which can lead to compromised security. Secure passwords might not be appropriate for use on mobile devices due to the length of time required for their input. The more individuals use their mobile device (and the more personal data is stored), the faster they'll want to be able to access that device. Many mobile devices provide PINs—sequences of numbers—as alternatives to passwords. PINs have the benefit of being able to be entered quickly, but they provide far less security than passwords and might be guessed more quickly.

Pattern locks provide a similar level of security to PIN codes; they allow users to choose a sequence of points during enrollment and repeat it during authentication. These are convenient and easy to remember—in some cases, users can authenticate without even looking at their phone's screen—but provide only a limited level of protection. Furthermore, pattern locks could be in full view of others with access to the phone's screen, and fingertip oils could leave a distinctive trace, indicating the pattern used to access the device.

We need a quick but secure authentication method for mobile devices. Biometrics could be the answer: they're present on the user at all times, they can't be forgotten, and they can be input as quickly as a glance or a touch.

Biometrics have the potential for additional capabilities on mobile devices. For example, certain biometric traits can indicate a person's stress level, so the system can use this information to respond in an appropriate manner. Mobile devices are personal and typically dedicated to an individual, making biometrics a natural choice for study and implementation.

Although biometrics provide many benefits, the security of the system's resources and the user's own biometric patterns must be considered.

Adopting biometrics on mobile devices could influence their use on desktop and other computing platforms. A challenge with using biometrics on traditional systems is providing better security than existing nonbiometric methods. However, as users become accustomed to the convenience of biometric authentication, expectations might be raised toward the adoption of biometrics elsewhere. Although nonmobile computing systems often provide additional resources, they're less portable. This might make the use of certain biometrics easier. For example, face recognition is more practical in a controlled, consistent environment. In other cases, some biometrics won't be possible or won't perform as well if the system can't move—for instance, tracking gait requires motion.

Although biometrics provide many benefits, the security of the system's resources and the user's own biometric patterns must be considered.

Biometrics

Biometrics can be physical or behavioral. Popular physical biometrics include fingerprint and face recognition. Physical biometrics are based on our intrinsic characteristics, which are often determined at birth or through random muscle movements over the course of our lives. Behavioral biometrics range from our typing pattern to our movements when we walk. Behavioral biometrics are the result of learned patterns of activity, although they're still shaped by our physical characteristics (for example, our voice or gait).

Unlike passwords, for which only an exact match is acceptable, biometrics typically seek a good (but not perfect) match for authentication. For example, in the case of fingerprint recognition, we would expect some variation due to perspiration, pressure (on the sensor), obstructions, and other factors. For this reason, a perfect match would be hard to attain. Instead, we seek a match that's close enough to provide a reasonable level of certainty that a user is who he or she claims to be. The system administrator can select the degree of similarity needed for a match. A high-throughput system that doesn't protect sensitive data could use a lower matching threshold to reduce the risk of incorrectly rejecting a user. A system protecting more valuable information might require a more precise match, even if legitimate users are accidentally denied access.

Fingerprint Recognition

Our fingertips contain patterns of ridges and valleys. These patterns form during gestation and are considered to be unique to individuals. Fingerprints increase the fingertip's surface area, making it more sensitive and providing better grip. Points at which ridges terminate or split (ridge endings and bifurcations, respectively) are known as *minutiae*. These minutiae, along with the associated ridge directions, make up the data most commonly used for fingerprint recognition: the template. More advanced sensors with higher resolutions can account for the width of the ridges and even the locations of pores on our fingertips.[2]

For authentication, individuals requesting access to a resource will present their fingertip to the system. The system must then compare the template generated from the new fingerprint to the template in its records. Depending on the degree to which the new fingerprint template matches that in the database and the acceptance parameters established by the system administrator, a user will be rejected or accepted. In certain instances, when the fingerprint quality appears to be poor, the system could immediately prompt the user to try again in hopes of obtaining a better sample.

The primary benefit of fingerprints is that they're closely tied to the individual, but this is also the main risk of using fingerprints for authentication. The consequences of users having their fingerprint copied by an attacker are severe. A fingerprint can't change: once a fingerprint is deemed invalid, a user can no longer use it as a form of authentication. Fingerprints don't have the same flexibility as text passwords, for which users are simply asked to change their password. This risk is compounded if the same fingerprints are used to authenticate multiple systems—known as *function creep*.

There are ways to use fingerprints more securely; however, we must use different techniques than those for traditional text passwords. We can protect text passwords with a one-way transformation. The same input will yield the same output, but the output shouldn't be able to be manipulated to reveal its formation. This works for text passwords because they require an exact match. Fingerprints require a close match, making a one-way hash impractical. One option is to encrypt the template data, but this still relies on key secrecy and decrypting template data to determine matches. Another alternative is the biometrics cryptosystem, wherein a fingerprint can be used to generate a cryptographic key.[3] The fingerprint template is discarded and only the cryptographic key is stored. A third option is template transformation, wherein users can supply an additional key used to generate a hash of their biometric information.[4]

Fingerprints are a useful and convenient biometric, but care must be taken when storing the sensitive data. Major smartphone vendors have been appropriately cautious by restricting access to the fingerprint sensor and taking steps to ensure that templates are protected and can't be removed from the device on which they were generated.

Biometrics are a promising method of authenticating users of mobile devices due to their convenience, but they must be used properly.

Facial Recognition

Facial recognition is the oldest and most popular biometric. It has special appeal on mobile devices, as camera hardware is relatively inexpensive and widely deployed on smartphones. Faces don't require physical contact to sample, and many users are accustomed to having their photo taken. Facial recognition functions similarly to fingerprint recognition, in the sense that the face is first sampled (often, a 2D picture is taken in the visible spectrum) and then transformed into a set of features that can be used for matching.[5]

There are challenges to implementing facial recognition reliably on mobile devices. Mobile devices, as their name suggests, can be used in any environment at any time and can be held at many different angles. This presents immense variability, compounded by the natural changes in our appearance in reaction to different weather, moods, stress, and other factors. Furthermore, we might change our appearance, perhaps by wearing glasses or growing a beard. It can be tempting to lower the acceptance threshold to help overcome the wide range of environmental factors that might interfere, but this could result in lower system security.

Facial recognition is easy to implement today, and major vendors have created libraries to simplify its deployment. However, practitioners must exercise caution and verify any solution before deployment. For example, *liveness detection* is particularly important in facial recognition. Liveness detection uses algorithms that can detect whether a sample is from a person who is truly in front of the device, or a potential replay attack, where the attacker attempts to record and reuse credentials that were previously validated. Without liveness detection, attackers can use a different image of a user (often easily found on the Internet) to gain access. Facial detection can provide convenient access, but this convenience should be weighed against security needs.

Behavioral Biometrics

Behavioral biometrics are patterns of human actions such as handwriting, voice, gait, and keystroke dynamics. One example of a behavioral biometric is the written signature, which has long been used to determine a document's authenticity. Behavioral biometrics are acquired and learned over time, although influenced by our physical characteristics, and are generally easy to collect. Often, sampling of behavioral biometrics can be done alongside another task without a user's direct involvement.

Behavioral biometrics are subject to great intrauser variation, meaning that the same metric can vary widely for a given user. Because of this, it's currently difficult to use behavioral biometrics alone for authentication, as a sufficiently low threshold of match acceptance must be used. Certain behavioral biometrics, such as voice, could be impersonated by others or subject to replay attacks.

Biometric authentication could also be combined with other methods of authenticating a system, such as initially requiring a lengthy (but secure) text password and then using biometrics for subsequent authentication attempts. If the quality of the biometric match decreases, the user can be prompted to reenter the text password instead.

Biometrics are a promising method of authenticating users of mobile devices due to their convenience, but they must be used properly. System security must be maintained and user biometrics must be protected. Biometrics are more usable than lengthy text passwords and are certainly better than no authentication at all, but their use must be tempered against security needs until biometrics authentication on mobile devices can be proven to be secure and reliable. ■

References

1. M. Harbach et al., "It's a Hard Lock Life: A Field Study of Smartphone (Un)Locking Behavior and Risk Perception," *Symp. Usable Privacy and Security* (SOUPS 14), 2014; www.usenix.org/system/files/conference/soups2014/soups14-paper-harbach.pdf.
2. D. Maltoni et al., *Handbook of Fingerprint Recognition*, Springer, 2005.
3. U. Uludag, S. Pankanti, and A.K. Jain, "Fuzzy Vault for Fingerprints," *Proc. 5th Int'l Conf. Audio- and Video-Based Biometric Person Authentication* (AVBPA 05), 2005, pp. 310–319.
4. A. Nagar, K. Nandakumar, and A.K. Jain, "Biometric Template Transformation: A Security Analysis," *Proc. SPIE7541, Media Forensics and Security II*, 2010, doi:10.1117/12.839976.
5. M.A. Turk and A.P. Pentland, "Face Recognition Using Eigenfaces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition* (CVPR 91), 1991, pp. 586–591.

Liam M. Mayron is a faculty member of the School of Computing, Informatics, and Decision Systems Engineering at Arizona State University. Contact him at lmayron@asu.edu.

Got an idea for a future article?
Email editors Richard Ford (rford@se.fit.edu) and Deborah A. Frincke (debfrincke@gmail.com).

This article originally appeared in IEEE Security & Privacy, *vol. 15, no. 3, 2015.*

Editor: **Liuqing Yang**, Colorado State University, liuqing.yang@ieee.org

Simple Is Beautiful: Toward Light Prediction Markets

Weiyun Chen, *East China Normal University*
Xin Li, *City University of Hong Kong*
Daniel D. Zeng, *University of Arizona*

A prediction market is a futures market in which participants trade contracts on the basis of probable outcomes of future events. Because individuals are rewarded for correct predictions, prediction markets encourage people to reveal private information, leading to accurate predictions of future events.[1] The first modern prediction market (Iowa Electronic Markets) was established more than 25 years ago.[2] Prediction markets have since attracted a lot of attention from researchers and practitioners. In this article, we focus on the cognitive load of prediction market mechanisms and argue for focused study on low-cognitive-load (that is, light) prediction markets.

As in a stock market, prediction market participants can buy or sell contracts on the market, in which they must assess the probabilities for each event outcome. Four steps make up a user's cognitive load for a particular event in the prediction market (see Figure 1). These steps are

- *timing*—when to purchase a contract;
- *pricing*—how much to pay for a contract;
- *revisiting*—whether to monitor the market after purchase; and
- *benefit*—whether the expected payoff is a determined value.

If the cognitive load of the four steps is too high and the economic and psychological return is not high enough, users' incentive to participate in future prediction market activities could be lowered.

In this article, we examine the cognitive load of the four steps of prediction markets and discuss how to design a prediction market mechanism with a low cognitive load.

Cognitive Load: Auction versus Posted Price Markets

Auctions, such as continuous double auctions (CDAs),[1] are commonly used to organize prediction markets. In a CDA-based prediction market, a user (that is, trader) can specify her intended contract price and submit a bid or ask request, which will be matched by the market platform with others' requests and form a transaction. The price for the bid or ask request directly reflects users' digestion of information related to the event.[3] The contract price formed on the market aggregates all related event information and can be used to predict event probability.

An alternative to an auction is the posted price market, in which one party reports her required price and waits for the other side to accept. In prediction markets, the posted price is often provided by a market maker (through an artificial trader). Prior studies have developed several market scoring rules (MSRs), that is, contract pricing algorithms, for market makers.[4] (In addition to their use in posted price markets, the MSRs can also be applied in CDA markets, where the artificial trader makes bid and/or call requests together with other users.)

From a user cognitive load perspective, auction-based mechanisms require users to specify their private price (indicating event probability),[5] which puts a high and continuous cognitive load on users in the pricing step of Figure 1.[6] However, a posted price market requires only that users make purchase decisions, which significantly lessens the load for users at this step.

Cognitive Load: Dynamic versus Deterministic Price and Payoff Markets

Two other features related to prediction market mechanisms are the contract price and payoff, which affect users' economic benefits. In CDAs, the contract price is dynamic and formed jointly by market participants' offers, whereas the payoff is often predetermined and known to participants from the beginning. The contract price reflects the

 Published by the IEEE Computer Society 2376-113X/15/$31.00 © 2015 IEEE

crowd's private information on future probabilities.

There also exist market mechanisms with fixed-price and dynamic payoffs, such as Pari-mutuel betting.[7] Pari-mutuel betting is a winner-takes-all mechanism in which all users buy contracts at the same price, and the correct predictors share the bid money in proportion to their bets. The payoff depends on the final bids on each side and is not predetermined. In this market, the final payoff odds reflect participants' estimated event probability with a favorite-longshot bias.[8]

Pari-mutuel betting can be combined with CDAs to form the Dynamic Pari-mutuel Mechanism (DPM).[9] DPM has an automatic market maker that sells the contract at a price determined by inventory, which changes nonlinearly with the amount of bets on each side. Its contract payoffs are determined by the total bids on each side when the market closes. Thus, DPM is a dynamic price and dynamic payoff mechanism.

From a cognitive load perspective, dynamic price increases users' cognitive load on deciding the timing to join the game and whether to revisit and monitor game progress after making purchases. Dynamic payoff affects users' certainty on payoff value and their information collection efforts.

Dynamic price mechanisms are also subject to market price manipulation.[10] As in financial markets, dishonest users can use their claimed price as a tool to distort the market price and earn economic benefits.[11] Furthermore, they can manipulate price to influence other users. For example, the 2008 US presidential election markets on Intrade.com were manipulated to urge more supporters to vote for one candidate.[12]

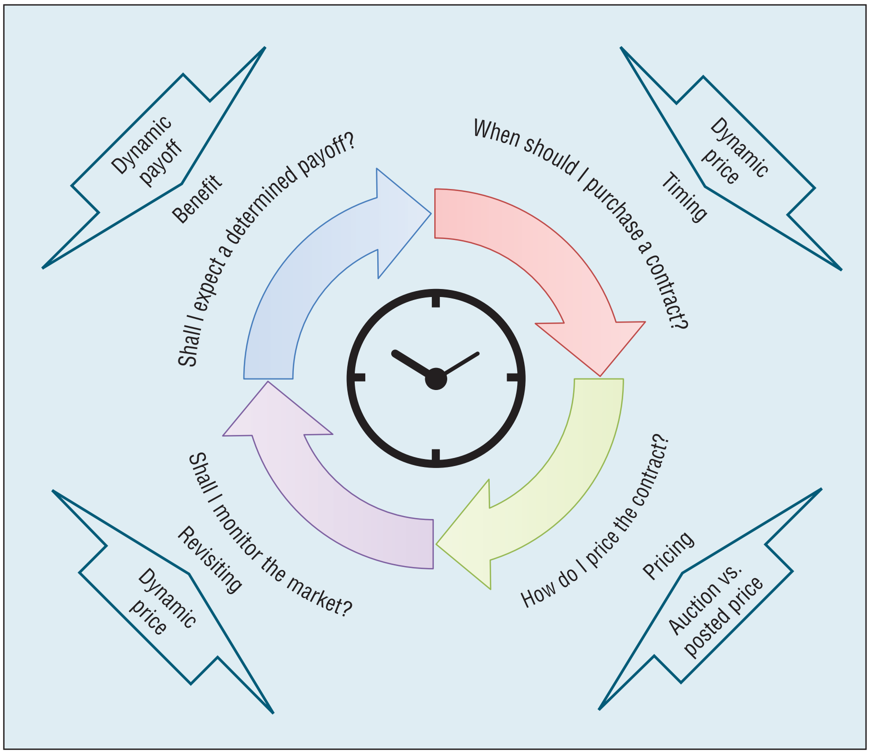


Figure 1. Four steps in the process of a prediction market game that affect a user's cognitive load, starting from timing, to pricing, revisiting, and benefit. The arrows in the outer circle represent prediction market features that could affect the cognitive load in each step.

Simple Is Beautiful

The combination of pricing mechanisms and price and payoff characteristics leads to multiple plausible prediction market mechanisms with different cognitive loads. Figure 2 illustrates some mechanisms with the three dimensions of mechanism design, highlighting the decision steps with high cognitive loads in a darker color for each mechanism. (Note that auctions cannot be fixed price, and those cells are not shown on the table.)

In this figure, the mechanisms to the right of the red curve have a relatively higher cognitive load. (We created a DPM without a CDA mechanism to represent using the pricing and payoff mechanisms of DPM in a posted price market. Such a mechanism does not seem to have been used before.) The mechanisms to the left of the red curve, such as Pari-mutuel, have a relatively lower cognitive load.

In a prediction market, cognitive load is a major participation cost for users. Given a stable economic stimuli, a high cognitive load could lead to a smaller number of users.[6] This is probably the reason that the high-cognitive-load DPM method has been used only in some experimental prediction markets. The low-cognitive-load Pari-mutuel method, however, has been widely employed in sports and entertainment prediction websites. For example, Goldderby.com uses the Pari-mutuel method to predict entertainment events, such as Oscar winners, and can often attract 500 users per prediction game.

In Figure 2, the MSR without CDA mechanism represents using the MSR for pricing in a posted price market. An example is the Microsoft prediction lab website launched in September 2014, which offers prices for users to accept. Its posted price nature

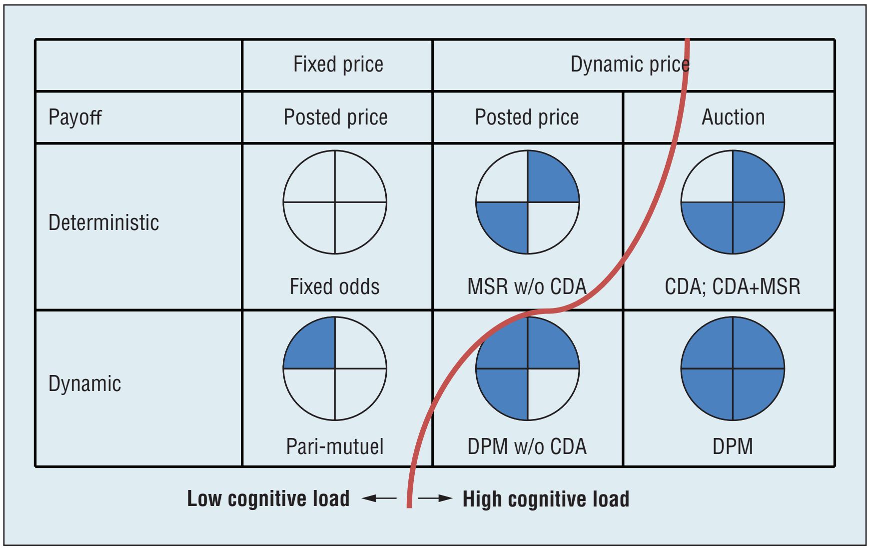


Figure 2. Cognitive load for different prediction market mechanisms. The pie chart shows the cognitive load for each of the four decision steps in Figure 1, where a darker color indicates a high cognitive load.

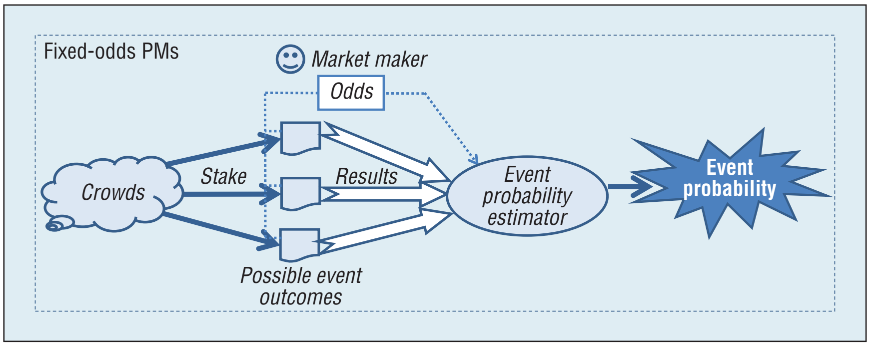


Figure 3. A generic fixed-odds prediction market architecture. The architecture combines an event probability estimator after a fixed-odds betting platform and makes predictions considering odds' impact on users.

attracted great interest from the media in discussing this "game-like polling technique." Although the market works on a classic topic—presidential election prediction—the reduced cognitive load from CDAs seemed to make the game more popular than its antecedents.

In this article, we argue that cognitive load is an important factor in concerns about participating in prediction markets. It will affect users' long-term incentive to stay in the prediction market and contribute their private information. Because prediction markets are based on individuals' private information, a simpler prediction market has a higher chance to sustain, especially in situations where more participants are necessary.

Fixed-Odds Prediction Market

In Figure 2, one cell reduces the cognitive load to a maximum extent: a posted price, fixed price, and determined payoff prediction market. This mechanism is essentially "fixed-odds betting" in gambling. According to our framework, it is a light market mechanism. However, can it really be used as a prediction market?

In fixed-odds betting, bettors (that is, users) buy contracts at a fixed price. Bookmakers (that is, market makers) set a fixed payoff (odds) before the betting starts. On the basis of the offered odds, users decide which side to bet to maximize their profit.[13] There is an obvious obstacle of using fixed-odds betting as a prediction market mechanism: its price and payoff are fixed, and it does not have a direct measure to capture the aggregated users' private information and report event probability. (If the odds are mispriced, the bookmaker will face significant economic loss.[13] Therefore, bookmakers usually strive to employ experts and sophisticated models to predict future event outcomes and to set odds. As a result, odds are correlated with future event probabilities. However, odds-based predictions are based on traditional modeling methods and do not reflect the crowd's intelligence.)

A Design for Fixed-Odds Prediction Markets

To address this concern, we suggest an architecture for a fixed-odds prediction market (see Figure 3). It has two major components: a betting platform and an event probability estimator. The betting platform is similar to a regular betting platform. When the betting stops, participants' assessments of event information are reflected in the bets on each side, which are influenced by the market odds. Such bets are fed into the event probability estimator $f()$ to decode an event probability from participants' stake:

$$P_{event} = f(\text{odds, bids}).$$

Building an event probability estimator for fixed-odds betting is not a trivial task. One approach is to model users' betting decisions. Each user's betting decision depends on the tradeoff of gain and loss, which is affected by the user's risk aversion function (on belief of an event's outcome) and odds (on return of an event's outcome). Higher belief and higher odds on an outcome will lead to higher expected utility. By assuming a relationship between crowd belief distribution and event probability, we can determine the most likely event probability that leads to a crowd belief distribution causing the final bet share of the game.

Under this framework, we employed regression on simulated data to approximate a reduced-form model on the crowds' mean belief, fixed odds, and the final betting ratio.[14] In recent studies, we built a structural model with crowd belief distribution as a linear combination of beta distributions and used maximum likelihood to estimate the parameters of the beta distributions.[15] They showed that with the help of such probability estimators, fixed-odds betting can achieve a comparable performance with auction-based prediction markets.

Expectations for a Fixed-Odds Prediction Market

If fixed-odds betting can be put into practice as a light prediction market mechanism, what should we expect?

- Simpler and more crowded auctions. As we've discussed, fixed-odds prediction markets have the lowest cognitive load among all prediction market mechanisms. The mechanism's simplicity lessens the load for users and could lead to a larger population of users and more information sources. The extra information has the benefit of more accurate predictions.
- Shorter auctions are better. In fixed-odds prediction markets, users have an incentive to bid later and wait for extra information, because the payoff is fixed. The odds set up by the market maker might also be more biased depending on how soon before the event they are set. Thus, a fixed-odds prediction market is better for a short-span game within which users' beliefs do not change much.

From these two conditions, we project most users of fixed-odds prediction markets will take a bet-and-leave strategy in playing the game,

We expect light prediction markets, such as those based on fixed-odds betting, to be an important branch of prediction markets.

which is different from the auction-based prediction markets that expect experienced users to keep updating their information through trading.

In short, we expect light prediction markets, such as those based on fixed-odds betting, to be an important branch of prediction markets. They will target the large population of less-experienced users involved in shorter games. The ongoing use of Pari-mutuel prediction markets and the Microsoft prediction lab shows the attractiveness of low-cognitive-load prediction markets. Ongoing research shows the feasibility of using fixed-odds betting as a prediction market mechanism. We project a great future for light prediction markets in research and practice.

Particularly, for fixed-odds prediction markets, there are several problems that need computer scientists and economists to work together to improve the mechanism design.

First, from a computer science perspective, it is necessary to explore more accurate, consistent, and efficient models to build the probability estimator for fixed-odds prediction markets. One possible direction is to model the dynamics of betting activities. We can also develop different models for different decision scenarios and applications.

Second, from an economics perspective, it is necessary to study the setup of odds in fixed-odds prediction markets. Previous literature on the odds of fixed-odds betting usually focused on the bookmaker's possible economic loss.[13] However, the goal of prediction markets is to acquire and aggregate information. It might reasonably operate at a financial loss for the benefit of gaining users' information.[16] It is also possible for fixed-odds betting to operate with virtual money (incentivizing participants by leaderboards), where financial loss is no longer a concern. In such scenarios, it is necessary to gauge the odd's impact on users' decisions and the final prediction, which will need the input of experimental economists.

Third, it is necessary to study the interaction between traditional and light prediction markets, which attract different segments of users. Experienced users can employ the information aggregated in light prediction markets to trade in a traditional prediction market. Such hierarchal information aggregation could lead to changes in user behaviors and prediction performance.

Finally, it would be interesting to investigate multistage fixed-odds betting. As we mentioned earlier, fixed-odds prediction markets better suit short-timespan games. In reality, we might want the game to be longer and attract more users. In such a scenario, it could be possible to set up a series of fixed-odds prediction markets that have related but different fixed odds. Such a setup will raise many design questions for computer scientists and economists.

Acknowledgments

This research is partially supported by NNSFC Grant #71025001 and City University of Hong Kong SRG 7004142 and 7004287.

References

1. J. Wolfers, and E. Zitzewitz, "Prediction Markets," *J. Economic Perspectives*, vol. 18, no. 2, 2004, pp. 107–126.
2. R. Forsythe et al., "Anatomy of an Experimental Political Stock Market," *Am. Economic Rev.*, vol. 82, no. 5, 1992, pp. 1142–1161.
3. D.S. Lee and E. Moretti, "Bayesian Learning and the Pricing of New Information: Evidence from Prediction Markets," *Am. Economic Rev.*, vol. 99, no. 2, 2009, pp. 330–336.
4. R. Hanson, "Combinatorial Information Market Design," *Information Systems Frontiers*, vol. 5, no. 1, 2003, pp. 107–119.
5. A. Madhavan, "Trading Mechanisms in Securities Markets," *J. Finance*, vol. 47, no. 2, 2012, pp. 607–641.
6. A. Graefe and J.S. Armstrong, "Comparing Face-to-Face Meetings, Nominal Groups, Delphi and Prediction Markets on an Estimation Task," *Int'l J. Forecasting*, vol. 27, no. 1, 2011, pp. 183–195.
7. R.E. Quandt, "Betting and Equilibrium," *Quarterly J. Economics*, vol. 101, no. 1, 1986, pp. 201–207.
8. P. Asch, B.G. Malkiel, and R.E. Quandt, "Racetrack Betting and Informed Behavior," *J. Financial Economics*, vol. 10, no. 2, 1982, pp. 187–194.
9. D.M. Pennock, "A Dynamic Pari-mutuel Market for Hedging, Wagering, and Information Aggregation," *Proc. 5th ACM Conf. Electronic Commerce*, 2004, pp. 170–179.
10. J. Hansen, C. Schmidt, and M. Strobel, "Manipulation in Political Stock Markets—Preconditions and Evidence," *Applied Economics Letters*, vol. 11, no. 7, 2004, pp. 459–463.
11. M. Blume, S. Luckner, and C. Weinhardt, "Fraud Detection in Play-Money Prediction Markets," *Information Systems and E-Business Management*, vol. 8, no. 4, 2010, pp. 395–413.
12. D. Rothschild and J. Wolfers, "Market Manipulation Muddies Election Outlook," *Wall Street J.*, 2 Oct. 2008; www.wsj.com/articles/SB122283114935193363.
13. T. Kuypers, "Information and Efficiency: An Empirical Study of a Fixed Odds Betting Market," *Applied Economics*, vol. 32, no. 11, 2000, pp. 1353–1363.
14. W. Chen, X. Li, and D.D. Zeng, "Exploring Crowds' Mean Belief in Fixed Odds Betting for Event Prediction," *Workshop Information Technologies and Systems*, 2012, pp. 217–222.
15. W. Chen, X. Li, and D.D. Zeng, "Modelling Fixed Odds Betting for Future Event Prediction," SSRN working paper, 20 Jan. 2015; http://ssrn.com/abstract=2552447.
16. Y.L. Chen and D.M. Pennock, "Designing Markets for Prediction," *AI Magazine*, vol. 31, no. 4, 2010, pp. 42–52.

Weiyun Chen is an assistant professor in the Department of Educational Information Technology at the East China Normal University. Contact him at weiyun.chen@ia.ac.cn.

Xin Li is an assistant professor in the Department of Information Systems of the College of Business at the City University of Hong Kong. Contact him at xin.li.phd@gmail.com.

Daniel D. Zeng is a professor in the Eller College of Management at the University of Arizona and an affiliated professor at the Institute of Automation at the Chinese Academy of Sciences. He is the corresponding author of this article. Contact him at zengdaniel@gmail.com.

This article originally appeared in IEEE Intelligent Systems, *vol. 30, no. 3, 2015.*

Internet of Things

Making the Hype a Reality

Arpan Pal, *Innovation Lab, Tata Consultancy Services*

The Internet democratized information. The Internet of Things (IoT) will democratize knowledge. As industrial researchers who work with businesses, we at Tata Consultancy Services see our customers in every industry looking for ways to create active knowledge and insight from IoT data. Organizations wish to innovatively apply this data to new business models in which understanding and addressing customer needs and demands is the focus (not just the product or service, as it is now), culminating in an improved B2C or B2B2C value chain for most businesses. Is IoT technology geared up to meet the challenge? What can we do to make the IoT hype and prospects a reality? Toward this end, I identify six key interest areas we should pay attention to—*s*calability, *p*rivacy, *a*ffordability, *c*ontext-awareness, *e*ase-of-development, and *s*ecurity, or *SPACES*. I elaborate on each of these here.

Scalability: Leverage Computing at the Edge

The IoT is all about putting sensors on physical objects and humans and connecting them to monitor, diagnose, or predict physical states and events. It's predicted that 40 to 50 billion IoT devices will be connected to the Internet by 2020 (www.cisco.com/web/about/ac50/ac207/crc_new/university/RFP/rfp13078.html). So, can the infrastructure handle the deluge of data that these devices will churn out? Probably not.

Network scalability is the traditional way to look at this problem. Scalability would be required in every layer of the network. At the physical layer, we would need cognitive radio[1] and TV white space reuse to handle the spectrum crunch.[2] In upper layers, there is IETF 6LoWPAN to connect devices to IP networks (http://datatracker.ietf.org/wg/6lowpan/documents/), IPv6 to handle scalability at the network layer (http://ipv6.com/articles/organizations/IETF-History-IPv6.htm), and the IETF's Constrained Application Protocol (www.ipso-alliance.org), MQTT (http://mqtt.org), and ZeroMQ (http://zeromq.org) at the application layer to provide lightweight data transport. Going forward, adaptive systems[3] on top of software-defined[4] and information-centric networks (https://irtf.org/icnrg) will also play an important role.

However, there is a more elegant way to handle the scalability issue—instead of passing sensor data through the Internet via edge devices such as routers and gateways, this data can be processed and analyzed at the edge itself, utilizing the compute power already available in edge devices. This new paradigm is known as *fog computing*.[5] Incorporating the fog infrastructure into the IoT is important to provide the compute scalability needed for the real-timeness and latency predictability demanded by the IoT applications. However, the heterogeneity of edge compute nodes and their unpredictable availability poses interesting research challenges,[6] and that's in addition to the usual scalability required for storage and management of huge amounts of data.

Privacy: Make It Measurable

All the hype around the IoT will come to nothing unless privacy issues associated with the collected sensor data are addressed to the satisfaction of all stakeholders. The Holy Grail of privacy is to have data that's both contextually useful and forever privacy-preserving. In the complex ecosystem of the IoT, this is never going to happen, because there is a need for data sharing across multiple entities. Can we have a platform that can automatically ensure conformance to privacy regulations (https://www.privacyrights.org/topics/11)? Additionally, we should also look into giving insight to data owners about the potential for privacy breaches of shared data vis-a-vis the utility they're getting from the IoT applications that use this data. Providing a "privacy metric"[7]

2376-113X/15/$31.00 © 2015 IEEE

to data owners that quantifies a potential privacy breach is one way to address this. However, the creation of such a metric depends on the data type and its intended use, and is hence quite complex to model.

Affordability: Crowdsourced Sensing

The return on investment for major IoT deployments isn't sufficiently motivating at this stage for most businesses. This is a primary reason why IoT applications have rarely moved beyond pilot deployments. There are deployment and cost issues when it comes to installing multitudes of sensing hardware across physical spaces. To make it more cost-effective, enterprises are looking into crowdsourcing sensing data from mobile phones where appropriate. Mobile phones are already pervasive, and they come with rich sensors such as cameras, microphones, accelerometers, gyroscopes, magnetometers, GPS, and altimeters. Participatory and opportunistic sensing using mobile phones will play a key role in IoT deployment.[8] A plethora of applications are possible using mobile phone sensors, prominent among them being road condition monitoring, driving behavior analysis,[9] traffic monitoring,[10] wellness/health monitoring, and so on.[11] In the future, robots and unmanned aerial vehicles carrying sensors will also contribute to this pervasive, affordable sensing paradigm.

Context-Awareness: What It Means for IoT Analytics

The IoT's real value will be realized only when we can derive physical contexts from the sensor data gathered. For sensors deployed on physical objects, this means answering questions such as what happened, where, when, and why. For sensors deployed on humans, the context means answering questions such as who is doing what, in where, or who is feeling what and why.[12] Thus, context-awareness in the IoT ultimately boils down to using analytics to detect a physical event, cause, identity, location, activity, physiology, and psychology from sensed data.

Ease of Development: Address Complexity

In the diverse IoT world, we can no longer afford to build vertical applications from scratch—a platform-based development approach promoting software re-use via APIs and services is a must.[13] Going further, one of the fundamental game-changers for successful IoT adoption will be the democratization of knowledge derived from IoT data. To make this happen, we need a crowdsourced application development and consumption ecosystem. Developing analytical applications for the IoT is a complex process that needs diverse knowledge of domains, sensors, algorithms, programming, and deployment infrastructure. A viable way to address such complexity is through a model-driven development (MDD) framework.[14] MDD is an approach that aims to model knowledge across different stakeholders (such as sensor providers, algorithm providers, domain experts, and infrastructure providers) by allowing separation of concerns for each stakeholder. It can assist an IoT application developer in easily creating an application based on data and goal descriptions.[15] Domain-agnostic semantic data interoperation[16,17] will also play a big role in such a framework.

Security: Why It's Different for the IoT

The security of anything IoT-related is one of the most talked-about concerns surrounding IoT applications. What does security mean in the IoT context? Is it any different or more complex than traditional IT systems? Fundamentally speaking, the answer is no. Security in IoT systems still means reliable data encryption before transmission and efficient management of the encryption key exchange. Of course, there is a scalability problem in distributing large numbers of keys across large numbers of devices. The resource constraints of sensing devices also add to this complexity. However, the real issue of IoT security isn't its implementation, but rather its implication. Because the IoT links the physical world to the cyberworld, any security breach in the cyberworld will affect the physical world and vice versa, leading to potentially catastrophic impacts.[18] Hence, we must look at IoT security in a more holistic manner, focusing on the minimization of such impacts.

The hype around IoT technology seems justified, given that it could disrupt many businesses. However, there are quite a few challenges that need to be addressed before the IoT becomes a reality and adds value to the business value chain. Quite a few end-to-end service integration platforms are being built to address one or more of the SPACES challenges I've discussed here (www.tcs.com/about/research/Pages/TCS-Connected-Universe-Platform.aspx).[19] Such platforms will play an extremely important role in the future successful deployment of IoT solutions.

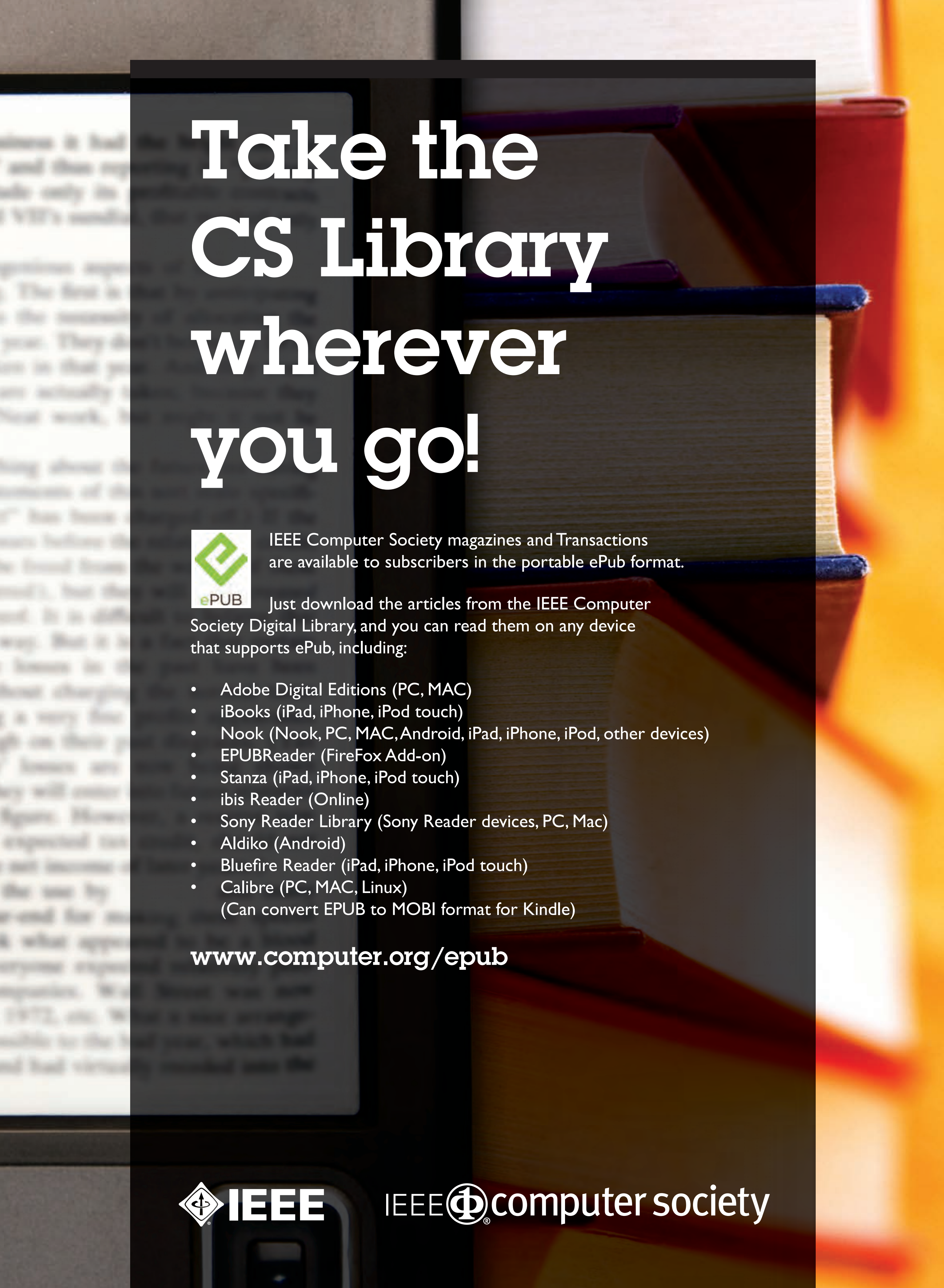
References

1. M.A. Shah et al., "Cognitive Radio Networks for Internet of Things: Applications, Challenges and Future," *Proc. 19th Int'l Conf. Automation and Computing* (ICAC), 2013, pp. 1–6.

This article originally appeared in
IT Professional, *vol. 17, no. 3, 2015.*

2. J. Temperton, "TV White Space Will Connect the Internet of Things," *Wired*, 13 Feb. 2015; www.wired.co.uk/news/archive/2015-02/13/white-space-spectrum.
3. S. Bandyopadhyay et al., "Adapting Protocol Characteristics of CoAP Using Sensed Indication for Vehicular Analytics," *Proc. 11th ACM Conf. Embedded Networked Sensor Systems*, 2013, article no. 48.
4. S. Fogarty, "7 Essentials of Software-Defined Networking," *Information Week Network Computing*, 19 Nov. 2013; www.networkcomputing.com/networking/7-essentials-of-software-defined-networking/d/d-id/898899.
5. B. Flavio et al., "Fog Computing and Its Role in the Internet of Things," *Proc. MCC workshop on Mobile Cloud Computing*, 2012, pp. 13–16.
6. S. Dey et al., "Challenges of Using Edge Devices in IoT Computation Grids," *Proc. Int'l Conf. Parallel and Distributed Systems* (ICPADS), 2013, pp. 564–569.
7. A. Ukil et al., "IoT-Privacy: To Be Private or Not to Be Private," *Proc. IEEE Conf. Computer Comm. Workshops*, 2014, pp. 123–124.
8. W.Z. Khan et al., "Mobile Phone Sensing Systems: A Survey," *IEEE Comm. Surveys & Tutorials*, vol. 15, no. 1, 2013, pp. 402–427.
9. A. Ghose et al., "Road Condition Monitoring and Alert Application: Using In-Vehicle Smartphone as Internet-Connected Sensor," *Proc. IEEE Int'l Conf. Pervasive Computing and Comm. Workshops*, 2012, pp. 489–491.
10. T. Chakravarty et al., "MobiDrive-Score—A System for Mobile Sensor-Based Driving Analysis: A Risk Assessment Model for Improving One's Driving," *Proc. 7th Int'l Conf. Sensing Technology* (ICST), 2013, pp. 338–344.
11. A. Pal et al., "A Robust Heart Rate Detection Using Smartphone Video," *Proc. 3rd ACM MobiHoc Workshop on Pervasive Wireless Healthcare*, 2013, pp. 43–48.
12. P. Charith et al., "Context Aware Computing for the Internet of Things: A Survey," *IEEE Comm. Surveys & Tutorials*, vol. 16, no. 1, 2013, pp. 414–454.
13. P. Balamuralidhara et al., "Software Platforms for Internet of Things and M2M," *J. Indian Inst. of Science*, vol. 93, no. 3, 2013, pp. 487–498.
14. P. Patel et al., "A Model-Driven Development Framework for Developing Sense-Compute-Control Applications," *Proc. 1st Int'l Workshop Modern Software Eng. Methods for Industrial Automation*, 2014, pp. 52–61.
15. A. Pal et al., "Model-Driven Development for Internet of Things: Towards Easing the Concerns of Application Developers," *Proc. Int'l Conf. IoT as a Service*, 2014.
16. *Semantic Sensor Network XG Final Report*, tech. report 28, W3C Incubator Group, June 2011.
17. D. Pfisterer et al., "SPITFIRE: Towards a Semantic Web of Things," *IEEE Comm. Magazine*, Nov. 2011, pp. 40–48.
18. E. Naone, "Hacking the Smart Grid," *Technology Rev.*, 2 Aug. 2010; www.technologyreview.com/news/420061/hacking-the-smart-grid/.
19. J. Gubbi et al., "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions," *Elsevier J. Future Generation Computer Systems*, vol. 29, 2013, pp. 1645–1660.

***Arpan Pal** is a principal scientist with Tata Consultancy Services, where he heads cyber-physical-systems research at the Innovation Lab, Kolkata. His current research interests include mobile phone and camera-based sensing and analytics, physiological sensing, M2M communications, and Internet-of-Things-based platforms and systems. Pal has more than 65 publications in reputed journals and conferences, has filed for more than 60 patents, and has nine patents granted to him. Contact him at arpan.pal@tcs.com.*


Take the CS Library wherever you go!
ePUB
IEEE Computer Society magazines and Transactions are available to subscribers in the portable ePub format.
Just download the articles from the IEEE Computer Society Digital Library, and you can read them on any device that supports ePub, including:
• Adobe Digital Editions (PC, MAC)
• iBooks (iPad, iPhone, iPod touch)
• Nook (Nook, PC, MAC, Android, iPad, iPhone, iPod, other devices)
• EPUBReader (FireFox Add-on)
• Stanza (iPad, iPhone, iPod touch)
• ibis Reader (Online)
• Sony Reader Library (Sony Reader devices, PC, Mac)
• Aldiko (Android)
• Bluefire Reader (iPad, iPhone, iPod touch)
• Calibre (PC, MAC, Linux)
(Can convert EPUB to MOBI format for Kindle)
www.computer.org/epub
IEEE
IEEE computer society

CAREER OPPORTUNITIES

MARKETING ANALYST, MEMPHIS, TN: Gather data, support, analyze, maintain marketing data/reports for IT resourcing; Work with vendors/end clients to gather data and prepare reports for management. Monitor marketing information/ reports continuously for qualified candidates/vendors/end clients. Reply to: SVS Technologies Limited, 8700 Trail Lake Drive, #228 Memphis, TN 38125

AUDITOR–TELECOMMUNICATION. Job Location: Charlotte, NC (HQ) with work-at-home option available for all cities through-out the U.S. Duties: Serve as secondary liaison for Telecom Expense Mgmt. (TEM) Business Services w/ accountability & reduce client's telecomm. costs; use NexTEM software to identify invoice & audit exceptions & follows standard remediation processes; & file billing disputes & change requests w/ telecomm. service providers, reconcile dispute filings, tariff. Reqs: US Bach. deg. or foreign equiv. in Telecomm. or rel. field plus 6 mths. of exp. in the job offered or as a Telecomm. Analyst or closely rel. field. Exp. which may have been obtained concurrently must incl. 6 mths. of exp. w/ NexTEM & tariff app. Please mail resume to: Jennifer Van Meter, Dimension Data North America, 11730 Plaza America Dr., Ste. 340, Reston, VA 20190

SIEMENS PLM SOFTWARE INC. has the following openings in Milford, OH: **Software Engineer Advanced**/Req#144526 to design, develop, modify & implement programming for products; **Software Engineer Sr.**/Req#144330 to perform design & software programming of products & services in File Mgmt area of Teamcenter product. Email resumes to PLMCareers@ugs.com & refer to Job code of interest. EOE

CLOUDERA, INC. is recruiting for our Palo Alto, CA office: Software Engineer: Plan, design, develop, test, document, & support new features, improvements to existing features & artifacts. Mail resume w/job code #36175 to: Cloudera, Attn.: HR, 1001 Page Mill Rd., Bldg. 2, Palo Alto, CA 94304.

MOBILEIRON, INC. is recruiting for our Mountain View, CA office: Sr. Software Engineer: Build, define and engineer user interface (UI) for server side components to support company's enterprise server product. Implement new core UI infrastructure using JavaScript, JSP, JQuery, CSS and HTML. Use object oriented programming languages like Java and JavaScript to incorporate fundamental design patterns such as MVC. Mail resume w/job code #35945 to: MobileIron, Attn.: HR, 415 E. Middlefield Rd., Mountain View, CA 94043.

SYSTEMS ADMINISTRATOR: Analyze, plan, design, install, configure administer, Implement and support s/w using knowledge in Sun Solaris 8/9/10.x, Redhat Linux, Veritas Volume Manager, HPUX 10.66, IBM WebSphere Application Server, EMC Symmetric, Clarion, Apache 2.x, Sun Studio, FileNet, NFS, Oracle, Sybase and MS SQL; Must be willing to travel & reloc; Reqs MS in comp sci, eng or rel. Mail resumes with Job id # RTSA0528 to Kaizen Technologies, Inc, 1 Lincoln Highway, Suite 10, Edison, NJ 08820.

SAMSUNG SEMICONDUCTOR INC. has a Sr. Engineer (job code: 5FE2508) job opportunity in San Jose, CA: Design, evaluate and specify circuit components, including VCO, DCO, inductor, capacitor bank and biasing circuit etc. Mail resume to Samsung Semiconductor, c/o Staffing – PTCL, 601 McCarthy Blvd., Milpitas, CA 95035. Must reference job code to be considered.

CLASSIFIED LINE AD SUBMISSION DETAILS: Rates are $425.00 per column inch ($640 minimum). Eight lines per column inch and average five typeset words per line. Send copy at least one month prior to publication date to: Debbie Sims, Classified Advertising, *Computer* Magazine, 10662 Los Vaqueros Circle, Los Alamitos, CA 90720; (714) 816-2138; fax (714) 821-4010. Email: dsims@computer.org.

In order to conform to the Age Discrimination in Employment Act and to discourage age discrimination, *Computer* may reject any advertisement containing any of these phrases or similar ones: "...recent college grads...," "...1–4 years maximum experience...," "...up to 5 years experience," or "...10 years maximum experience." *Computer* reserves the right to append to any advertisement without specific notice to the advertiser. Experience ranges are suggested minimum requirements, not maximums. *Computer* assumes that since advertisers have been notified of this policy in advance, they agree that any experience requirements, whether stated as ranges or otherwise, will be construed by the reader as minimum requirements only. *Computer* encourages employers to offer salaries that are competitive, but occasionally a salary may be offered that is significantly below currently acceptable levels. In such cases the reader may wish to inquire of the employer whether extenuating circumstances apply.

ADVERTISER INFORMATION • JULY 2015

Advertising Personnel

Debbie Sims: Advertising Coordinator
Email: dsims@computer.org
Phone: +1 714 816 2138 | Fax: +1 714 821 4010

Chris Ruoff: Senior Sales Manager
Email: cruoff@computer.org
Phone: +1 714 816 2168 | Fax: +1 714 821 4010

Advertising Sales Representatives (display)

Central, Northwest, Far East:
Eric Kincaid
Email: e.kincaid@computer.org
Phone: +1 214 673 3742
Fax: +1 888 886 8599

Northeast, Midwest, Europe, Middle East:
David Schissler
Email: d.schissler@computer.org
Phone: +1 508 394 4026
Fax: +1 508 394 1707

Southwest, California:
Mike Hughes
Email: mikehughes@computer.org
Phone: +1 805 529 6790

Southeast:
Heather Buonadies
Email: h.buonadies@computer.org
Phone: +1 201 887 1703

Advertising Sales Representatives (Classifieds & Jobs Board)

Heather Buonadies
Email: h.buonadies@computer.org
Phone: +1 201 887 1703

ATTN: J. Aldax, 250 Brannan Street, San Francisco CA 94107. Individuals seeking employment at Splunk are considered without regards to race, religion, color, national origin, ancestry, sex, gender, gender identity, gender expression, sexual orientation, marital status, age, physical or mental disability or medical condition (except where physical fitness is a valid occupational qualification), genetic information, veteran status, or any other consideration made unlawful by federal, state or local laws. To review US DOL's EEO is The Law notice please visit: https://careers.jobvite.com/Splunk/EEO_poster.pdf. To review Splunk's EEO Policy Statement please visit: http://careers.jobvite.com/Careers/Splunk/EEO-Policy-Statement.pdf. Pursuant to the San Francisco Fair Chance Ordinance, we will consider for employment qualified applicants with arrest and conviction records.

COMPUTERS: Tech Mahindra (Americas) Inc. is seeking to fill numerous IT positions. Prgrm Mgrs to oversee & manage mult. IT projects, proj. planning, dvlpmnt, implementation, acct & delivery mgmt; Proj. mgrs to oversee & manage IT teams w/dvlpmnt of various sftwre apps. Sys. Analyst/Programmers/ Quality & Tech Architect/ Tech Solutions Architect / Quality Engg /Test Eng/ Sftwre Eng. / sftwre Developer / Systems Admin/DBAs/ DB Architect to analyze, design, dvlp, test & maintain comp software apps or databases through all phases of sftwre dvlpmt life cycle (Sftwre Eng. may also lead a team on various projects); Telecom solutions Architect / Application Architects to use various technologies to dsgn & dvlop telecom software apps. Mech. Eng (CAD/CAM) to design, develop, validate & perform structural calculations, product improvement, & provide tech. support to design teams at high levels utilizing specific mech. tools. Sales Eng/ Bus. Analyst/ Mgmt Analyst for solutions/ pre-sales activities w/relev. industry experience. IT Bus. Dev. Mgrs to create new business, negotiate contracts & dvlp proposals for customized IT solutions. Rel. Mgrs to manage/outsource commercial IT/Eng. deals, monitor & maintain existing accts. All Tech/Mgrial. positions require a MS or BS degree or equiv. in CS, Comp Apps, CIS, IT, Eng, Bus mgmt/admin or closely related fields and relevant industry exp. Sales/Rel.Mgrs. require a MS or BS degree or equiv. in Bus. Admin, Eng. or closely related field & relev. Industry exp. Positions are based out of corp. HQ in 4965 Preston Park Blvd, # 500, Plano, TX 75093 & subject to travel & relocation to client sites located throughout the U.S. Mail resume & position applied for w/JOB CODE: 07IE15 to RMG, Tech Mahindra (Americas) Inc., 1001 Durham Avenue, Suite 101, South Plainfield NJ 07080.

JUNIPER NETWORKS

Juniper Networks is recruiting for our Sunnyvale, CA office:

Technical Support Engineer Staff #35300: Provide high level technical support/guidance on company routing products to customers. Deliver in-depth diagnostics and root-cause analysis for network-impacting issues on company routing products (Internet backbone routers) to large Internet Service Provider and/or enterprise customers.

Software Engineer Staff #7048: Design, implement, test and debug security features in flow based packet forwarding engine of company's NGSRX platforms. Communicate with product line managers (PLM) to develop functional specifications of features.

Hardware Engineer #32827: Perform analysis and design and understand tradeoffs in designing interconnect solutions ranging from chip to chip, board to board, backplane and chassis to chassis interconnect. Perform channel margin analysis to provide design tradeoffs amongst package, board, and connector.

Software Engineer #12183: Design, develop, and enhance existing and new services on company routers, including NAPT ALGs, Carrier Grade NAT (CCGNAT), network stack enhancements and platform infrastructure. Address performance and scalability requirements and issues during the design and development of features.

Technical Support Engineer #29771: Work with customers to resolve technical and non-technical problems related to routers, protocols and network design. Troubleshoot complicated hardware and software issues, and replicate customer environments and network problems in the lab.

ASIC Engineer #18214: Develop, specify and architect the next generation of Fabric ASICs. Use the Verilog Hardware Description Language to design digital logic, based on architectural specification.

Product Manager #31146: Drive products through the product lifecycle and develop product roadmap. Collaborate with engineering, sales, and manufacturing teams to product manage one of the core routing product families at Juniper.

Technical Support Engineer #5699: Deliver in-depth diagnostics and root-cause analysis for network impacting issues on Juniper routing products to large ISPs and/or enterprise customers.

ASIC Engineer #30172: Develop test plan, test bench for cutting edge technology networking chips. Write test objects and carry on tests. Perform post-silicon validation in the lab.

Software Engineer #29958: Design and develop networking software for data center, campus and service provider networking equipment. Triage and resolve customer issues reported on networking software.

QA Engineer #23508: Define and create test plans and test cases based on system requirements and specifications. Create automation scripts for test cases. Execute manual testing and automated scripts and verify the results.

Test Engineer #34200: Create solid test plans and procedures, and enhance test plans with customer deployment scenarios. Perform test executions and completions on time across releases.

Systems Engineering Sr. Manager #8641: Provide leadership and guidance to a team of Systems Engineers to help design, evaluate and deploy Data networking solutions within Large Enterprise Data Centers and Web hosting providers. Telecommuting allowed and occasional travel required.

Test Engineer Staff #8192: Test complex multi-service networking and network security products such as L4-L7 security infrastructure, SSL proxy, web security, application security, and application service components.

Juniper Networks is recruiting for our San Francisco, CA office:

Security Engineer Staff #32223: Design and code cloud-based security services. Work alongside technical leads on critical aspects of the security software design and implementation.

Juniper Networks is recruiting for our Westford, MA office:

Test Engineer #26245: Design, develop, and document system level verification tests for Juniper routers, and work with software and hardware teams to develop product verification, high availability, and scalability tests for network routers.

Technical Support Engineer #30743: Work with customers and resolve issues related to network gear. Follow up and communicate with customers on progress and the status of their open case and user experience.

Mail single-sided resume with
job code # to
Juniper Networks
Attn: MS 1.4.251
1133 Innovation Way
Sunnyvale, CA 94089

Software Development Engineer (Req#9RGSDC). Dev SW to enable the testing of high-performance, large scale computing systems.

Software Development Engineer (Req#9F23H9). Des & dev kernel SW, sys SW, & tools for SW debug & performance analysis.

Hardware Development Engineer (Req#9B2TE9). Respon for design, develop, characterization & test sensor sys for consumer elect prods.

Software Engineer Applications (Req#9C4U6U). Conduct performance analysis of Mac platforms.

Software Development Engineer (Req#9DPM5W). Design and develop cellular SW features.

Hardware Development Engineer (Req#9E3V62). Res for all asp'ts of acoustic des for vari audio accessory prods. Travel req: 20%.

Software Engineer Applications(Req# 9F222P). Des & dev real time rendering algo'ms on variety of HW plat's.

Senior Hardware Development Engineer (Req#9EZ3AS). Des, dev & vald radio freq. HW circuits & syst for wireless comm devices.

Senior Platform Process Engineer (Req#9P3TRD). Manufac research & dev to solve complex probls that directly impact future prod, qual & delivery of Apple Displays. Travel Required: 15%.

System Design Manager (Req# 9E5VAG). Mng team of syst design engineers involved in evaluating HW syst. Travel Required: 25%.

Hardware Development Engineer (Req#9GW4HX). Define HW Equipment and Process baseline. Travel req'd 25%.

Software Development Engineer (multiple positions open) (Req# 9JX3DK). Resp for the bring-up and devlpmnt of next generation Macintosh platforms.

Senior Software Engineer (Req# 9TEVBL). Build high perform, high scalable, fault tolerant back-ends for critical internal sys.

Software Development Engineer (Req#9H3UG7). Design, develop, support & test graphics compilation sw.

Software Development Engineer (Req#9GANB3). Perform use case analysis to select graphics processing unit (GPU) workloads of spec interest.

Software Engineer Applications (Req#9RFVLD). Des & dev SW-based solutions for sales eng teams.

Apple Inc. has the following job opportunity in Jersey City, NJ:

Systems Design Engineer (Req# 9HDNDP). Test cellular telephony functionality of iOS devices. Travel req'd 30%.

Refer to Req#
& mail resume to
Apple Inc., ATTN: L.M.
1 Infinite Loop 104-1GM
Cupertino, CA 95014.

Apple is an EOE/AA m/f/disability/vets.

IT PROFESSIONALS. Sagitec Solutions, LLC, an established global technology solutions company seeks IT Professionals/Managers: Deputy Project Managers, Senior Software Engineers, Software Engineers, & Programmer Analysts. Deputy Project Manager reqs Master's or equiv. in Engg (any), CS, IT or related; 12 mos' relevant indus exp.; & Scrum Master certification. Senior Software Engineer reqs Master's or equiv. in CS, IT Engg or related & 12 mos' relevant indus exp. For both positions, Sagitec will also consider applicants w/ Bach's or equiv. in the stated fields & 5 yrs' progressively responsible & relevant indus exp. Software Engineer reqs bach's or equiv. in the stated fields & 24 mos' relevant indus exp. Positions based out of Sagitec HQ in Little Canada, MN & subject to relocation to various unanticipated sites throughout the U.S. Software Engineer reqs Bach's or equiv. in CS, IT, Engg or related & 24 mos' relevant indus exp. Position based in Englewood, CO. Programmer Analyst reqs Bach's in CS, Engg (any) or related & 24 mos' relevant indus exp. Pension & retirement systems domain exp. req'd. Position based in Topeka, KS. Senior Software Engineer reqs Master's or equiv. in CS, IT Engg or related, & 12 mos' relevant indus exp. or Bach's or equiv. in the stated fields & 5 yrs' progressively responsible & relevant indus exp. Position based in Frankfort, KY or Topeka, KS depending on project needs. Mail resumes to: Sagitec Solutions, LLC, ATTN: HR, 422 County Road D. East, Little Canada, MN 55117.

APPLICATIONS SYSTEMS ANALYST/PROGRAMMER. For Roanoke, VA employment with TMEIC International Corporation. Develop Windows & web based software applications. Design, implement, test, deploy & document software systems to configure, price & output quotations for industrial motors & drives. Identify & define interfaces to other software applications & functions, including export management software. Plan & execute productivity enhancements or product development projects. Resolve internal &/or external customer issues with the developed software. Requirements: Bachelor's in Electrical Engineering or Computer Science or a related field. 3 yrs experience developing Windows rich-client applications using C# & Windows Forms. 3 yrs experience in the design & development of: CPQ applications for industrial motors & drives for the oil & gas, mining or testing industries; & SOA & web services integrating Oracle R12 with the Tradesphere & Oracle CRM applications. Experience: with WCF, MVC & EAI; & designing & implementing web services. Apply to Job # IT15018 at http://www.tmeic.com/North%20America/TWprMg.

(HCL422). **Project Managers:** Responsible for managing, planning, coordinating, supervising & directing IT professionals. **Cary, NC:** Reqs MS+1/BS+5 (HCL423). **Frisco, TX:** Reqs MS+1/BS+5 (HCL424). **Redmond, WA:** Reqs MS+1/BS+5 (HCL425). **Sunnyvale, CA:** Reqs MS+1/BS+5 (HCL426). **Jersey City, NJ:** Reqs MS+1/BS+5 (HCL427). * Employer will accept a combination of education & experience as determined by a qualified evaluation service as equivalent to a Bachelor's degree. Multiple job openings are available. How to apply: Mail resume, referencing HCL job code, including job history, to: HCL America, Inc., Attn: ISG Department, 330 Potrero Ave, Sunnyvale, CA 94085. HCL is an Equal Opportunity Employer.

HCL AMERICA, INC. has the following openings in multiple locations. Travel & work @ various unanticipated client sites as assigned: **Sales Engineers:** Responsible for selling various IT services to clients & pre-sales engineering support & guidance to customers. **Cary, NC:** Reqs BS*+0 (HCL428); BS+2 (HCL429); MS+1/BS+5 (HCL430). **Frisco, TX:** Reqs BS*+0 (HCL431); BS+2 (HCL432); MS+1/BS+5 (HCL433). **Redmond, WA:** Reqs BS*+0 (HCL434); BS+2 (HCL435); MS+1/BS+5 (HCL436). **Sunnyvale, CA:** Reqs BS*+0 (HCL437); BS+2 (HCL438); MS+1/BS+5 (HCL439). **Jersey City, NJ:** Reqs BS*+0 (HCL440); BS+2 (HCL441); MS+1/BS+5 (HCL442). **Business Development Managers:** Responsible for business development, discussions, planning, & product development, & working towards connecting & building relationships with key decision makers in customer organizations. **Cary, NC:** Reqs MS+1/BS+5 (HCL443). **Frisco, TX:** Reqs MS+1/BS+5 (HCL444). **Redmond, WA:** Reqs MS+1/BS+5 (HCL445). **Sunnyvale, CA:** Reqs MS+1/BS+5 (HCL446). **Jersey City, NJ:** Reqs MS+1/BS+5 (HCL447). **General & Operations Managers:** Coordinate business initiatives & integrate people processes across the company, & spearhead development, communication, & implementation of effective growth strategies & processes. **Cary, NC:** Reqs MS+1/BS+5 (HCL448). **Frisco, TX:** Reqs MS+1/BS+5 (HCL449). **Redmond, WA:** Reqs MS+1/BS+5 (HCL450). **Sunnyvale, CA:** Reqs MS+1/BS+5 (HCL451). **Jersey City, NJ:** Reqs MS+1/BS+5 (HCL452). **Sales Managers:** Responsible for sales operations in the area of infrastructure management, managing & directing sales operations, & carrying out business development. **Cary, NC:** Reqs MS+1/BS+5 (HCL453). **Frisco, TX:** Reqs MS+1/BS+5 (HCL454). **Redmond, WA:** Reqs MS+1/BS+5 (HCL455). **Sunnyvale, CA:** Reqs MS+1/BS+5 (HCL456). **Jersey City, NJ:** Reqs MS+1/BS+5 (HCL457). **Senior Market Research Analysts:** Conduct market research & price comparisons, & analyze market data to identify the competition, target groups, & market opportunities. **Frisco, TX:** Reqs MS+1/BS+5 (HCL458). *Employer will accept a combination of education and experience as determined by a qualified evaluation service as equivalent to a Bachelor's degree. Multiple job openings are available. How to apply: Mail resume, referencing HCL job code, including job history, to: HCL America, Inc., Attn: ISG Department, 330 Potrero Ave, Sunnyvale, CA 94085. HCL is an Equal Opportunity Employer.

WHATSAPP, INC. currently has the following openings in Mountain View, CA (various levels/types): Software Engineer (2797) Help build the next generation of systems behind WhatsApp's products, create web/and or mobile applications that reach over one billion people, and build high-volume servers to support our content. Mail resume to: WhatsApp, Inc. c/o Facebook Inc. Attn: SB-GMI, 1 Hacker Way, Menlo Park, CA 94025. Must reference job title and job# shown above, when applying.

Help build the next generation of systems behind Facebook's products.

Facebook, Inc.

currently has the following openings in **Menlo Park, CA** (various levels/types):

Product Designer (5902J) Design, prototype, & build new features for Facebook's website or mobile applications. **Infrastructure Strategy Analyst (4487J)** Perform industry, market, & financial analysis related to the growth and improvement of Facebook's Infrastructure systems & computer applications. **Developer Support Engineer (2969J)** Interface with operations & engineering teams to drive development & improvement of application tools & processes. **Engineering Manager (819J)** Drive engineering effort, communicate cross-functionality, & be a subject matter expert; &/or perform technical engineering duties & oversee a team of engineers. **Program Manager, Education (2285J)** Play a key role in recruiting & building strategic relationships with relevant educators & education communities both internally at Facebook & externally at school campuses & within education groups.

Mail resume to: Facebook, Inc. Attn: SB-GMI, 1 Hacker Way, Menlo Park, CA 94025. Must reference job title and job# shown above, when applying.

Samsung Research America, Inc.

has the following opportunities (various levels) available in **Cambridge, MA**:

Senior Software Engineer (Job code: CBMA15J01)

Software Engineer, Staff 1 (Job code: CBMA15J02)

Specific requirements apply. All of these positions will involve developing technologies for company's computer, digital television, mobile telephone, printer, or other electronic products.

Mail your resume referencing job title and job code to farhat.k@samsung.com.

Linkedin Corp.

has openings in our **Mtn View, CA** location for:

Software Engineer (All Levels/Types) (6597.1052, 6597.1132, 6597.659, 6597.1160, 6597.1239, 6597.664, 6597.836, 6597.706, 6597.850, 6597.894, 6597.840, 6597.987, 6597.761, 6597.730, 6597.893, 6597.403, 6597.1219, 6597.1293, 6597.796, 6597.741, 6597.1003, 6597.739, 6597.663, 6597.976, 6597.1172, 6597.946) Design, develop & integrate cutting-edge software technologies; **Senior Site Reliability Engineer (6597.734)** Serve as a primary point responsible for the overall health, performance, & capacity of one or more Internet-facing service; **User Experience Designer (6597.829)** Design creative solutions that address business, brand & user experience requirements; **Test Engineer (6597.830)** Responsible for driving test automation for LinkedIn products through web & mobile UI automation; **Test Engineer (6597.947)** Drive test automation of products, working within a highly distributed platform capable of scaling beyond 340 million members, serving billions of page views; **Technical Program Manager, Security Ecosystem (6597.1148)** Develop a security researcher recognition program, & conduct outreach to the security community, including identifying thought leadership opportunities for Security team; **Associate Network Operations Center Engineer (6597.534)** Monitor & resolve application, system, & network incidents affecting the company platform & ensure maximum availability; **Linux Systems Engineer (6597.912)** Participate in operational duties for the Compute Operations team, including management of existing systems & deployment of new ones; **Senior Site Reliability Engineer (6597.732)** Serve as a primary point responsible the overall health, performance, & capacity of our internet-facing systems; **Test Engineer (6597.891)** Write & build automated suites, continuously design creative ways to break software, & identify potential bugs; **Senior Data Scientist (6597.866)** Extract & analyze data to drive product strategy; **Technical Program Manager (6597.940)** Manage the development & implementation process of products & services.

Linkedin Corp. has openings in our **Sunnyvale, CA** location for:

Software Engineer (All Levels/Types) (6597.473, 6597.1188, 6597.1114, 6597.879, 6597.714, 6597.1228, 6597.1232, 6597.993, 6597.1248, 6597.1271, 6597.1033, 6597.1152) Design, develop & integrate cutting-edge software technologies; **Computer Systems Analyst (6597.839)** Responsible for the analysis, design, & development of sales systems projects in conjunction with other members of **Sales Systems teams; Data Scientist, Business Insights (6597.911)** Improve the quality of data for high fidelity metrics, with a focus on bot detection; **Engineering Manager, Site Reliability (6597.311)** Lead a team of more than 3 engineers responsible for the operational design & health of revenue generating (jobs, ads, subscriptions) & mobile front-end systems; **Test Engineer (6597.852)** Design, develop & integrate cutting-edge software technologies.

Linkedin Corp. has openings in our **San Francisco, CA** location for:

Software Engineer (All Levels/Types) (6597.808, 6597.847, 6597.827) Design, develop & integrate cutting-edge software technologies; **Senior UI Engineer (6597.1344)** Lead development of the presentation & interactivity layer of web properties, including the creation of web pages & the UI of web-based tools; **Test Engineer (6597.1018)** Design & develop advanced test suites using object-oriented methodologies; **Solutions Analyst (6597.926)** Utilize frameworks for structuring problem solving, performing analysis, & developing recommendations.

Linkedin Corp. has openings in our **New York, NY** location for:

Software Engineer (All Levels/Types) (6597.1063) Design, develop & integrate cutting-edge software technologies;

Linkedin Corp. has openings in our **Calabasas, CA** location for:

Quality Assurance Engineer (6597.1346) Test user applications & features.

Please email resume to: 6597@linkedin.com. Must ref. job code above when applying.

Focus on
Your Job Search
IEEE Computer Society Jobs helps you easily find a new job in IT, software development, computer engineering, research, programming, architecture, cloud computing, consulting, databases, and many other computer-related areas.
New feature: Find jobs recommending or requiring the IEEE CS CSDA or CSDP certifications!
Visit www.computer.org/jobs to search technical job openings, plus internships, from employers worldwide.
http://www.computer.org/jobs
IEEE computer society | JOBS
The IEEE Computer Society is a partner in the AIP Career Network, a collection of online job sites for scientists, engineers, and computing professionals. Other partners include Physics Today, the American Association of Physicists in Medicine (AAPM), American Association of Physics Teachers (AAPT), American Physical Society (APS), AVS Science and Technology, and the Society of Physics Students (SPS) and Sigma Pi Sigma.

IEEE Cloud Computing Call for Papers

Although cloud technologies have been advanced and adopted at an astonishing pace, much work remains. *IEEE Cloud Computing* seeks to foster the evolution of cloud computing and provide a forum for reporting original research, exchanging experiences, and developing best practices.

IEEE Cloud Computing magazine seeks accessible, useful papers on the latest peer-reviewed developments in cloud computing. Topics include, but aren't limited to:

- Cloud architectures (delivery models and deployments),
- Cloud management (balancing automation and robustness with monitoring and maintenance),
- Cloud security and privacy (issues stemming from technology, process and governance, international law, and legal frameworks),
- Cloud services (cloud services drive and are driven by consumer demand; as markets change, so do the types of services being offered),
- Cloud experiences and adoption (deployment scenarios and consumer expectations),
- Cloud and adjacent technology trends (exploring trends in the market and impacts on and influences of cloud computing),
- Cloud economics (direct and indirect costs of cloud computing on the consumer; sustainable models for providers),
- Cloud standardization and compliance (facilitating the standardization of cloud tech and test suites for compliance), and
- Cloud governance (transparency of processes, legal frameworks, and consumer monitoring and reporting).

Submissions will be subject to *IEEE Cloud Computing* magazine's peer-review process. Articles should be at most 6,000 words, with a maximum of 15 references, and should be understandable to a broad audience of people interested in cloud computing, big data, and related application areas. The writing style should be down to earth, practical, and original.

All accepted articles will be edited according to the IEEE Computer Society style guide. Submit your papers through Manuscript Central at https://mc.manuscriptcentral.com/ccm-cs.

If you have any questions, feel free to email lead editor Brian Kirk at bkirk@computer.org.

www.computer.org/cloudcomputing

IEEE LCN
Local Computer Networks Conference

LCN 2015:

40th Annual IEEE Conference on Local Computer Networks

Call for Participation and Celebration

26 October - 29 October 2015
Sheraton Sand Key Resort
Clearwater Beach, Florida, USA

The IEEE LCN conference is celebrating 40 years as the premier conference on theoretical and practical aspects of computer networking. LCN is highly interactive, enabling an effective interchange of results and ideas among researchers, users, and product developers. Major developments from high-speed networks to the global Internet to specialized sensor networks have been reported at past LCNs. Please join us as we celebrate 40 years of LCN with a look back on the contributions made by past speakers, attendees, and authors.

Visit http://www.ieeelcn.org for more information.

Topics of interest include, but are not limited to:

- Personal and wearable networks
- Wireless ad hoc, sensor & RFID networks
- Information-centric networking
- Embedded networks
- Opportunistic networking
- Delay-tolerant networks
- Cognitive radio networks
- Vehicular and underwater networks
- Social networks
- Mobile and ubiquitous networking
- Green networking
- Overlay and peer-to-peer networks
- Local-area and metropolitan-area networks
- Storage-area networks
- Optical and high-speed access networks
- Software Defined Networking
- Internet of Things
- Link technologies
- Adaptive networking applications
- Authentication, authorization, accounting
- Security and privacy
- Cross-layer optimization
- Mobility and Location-dependent services
- Multimedia and real-time communication
- Machine-to-machine communications for smart environments
- Network management, reliability, and QoS
- Network traffic characterization and measurements
- Performance evaluation of networks
- Testbeds for network experiments
- Network coding

Don't miss the opportunity to be part of the 40th anniversary!

IEEE TCCC Technical Committee on Computer Communications IEEE computer society