COMPUTING

> Careers

1

1

00

- > Artificial Intelligence
- > Internet
- > Social Media

OCTOBER 2019

e

1

0

01

1

www.computer.org

0

1



Keep Your Career Options Open

Upload Your Resume Today!

TEMPLATES

RESUMES VIEWED

BY TOP EMPLOYERS

Whether your enjoy your current position or you are ready for change, the IEEE Computer Society Jobs Board is a valuable resource tool.

Take advantage of these special resources for job seekers:



No matter your career level, the IEEE Computer Society Jobs Board keeps you connected to workplace trends and exciting new career prospects.

www.computer.org/jobs





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







STAFF

Editor Cathy Martin

Publications Operations Project Specialist Christine Anthony

Production & Design Carmen Flores-Garvey Publications Portfolio Managers Carrie Clark, Kimberly Sperka

Publisher Robin Baldwin

Senior Advertising Coordinator Debbie Sims

Circulation: ComputingEdge (ISSN 2469-7087) 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 address changes to ComputingEdge-IEEE Membership Processing Dept., 445 Hoes Lane, Piscataway, NJ 08855. Periodicals Postage Paid at New York, New York, and at additional mailing offices. Printed in USA.

Editorial: Unless otherwise stated, bylined articles, as well as product and service descriptions, reflect the author's or firm's opinion. Inclusion in ComputingEdge does not necessarily constitute endorsement by the IEEE or the Computer Society. All submissions are subject to editing for style, clarity, and space.

Reuse Rights and Reprint Permissions: Educational or personal use of this material is permitted without fee, provided such use: 1) is not made for profit; 2) includes this notice and a full citation to the original work on the first page of the copy; and 3) does not imply IEEE endorsement of any third-party products or services. Authors and their companies are permitted to post the accepted version of IEEE-copyrighted material on their own Web servers without permission, provided that the IEEE copyright notice and a full citation to the original work appear on the first screen of the posted copy. An accepted manuscript is a version which has been revised by the author to incorporate review suggestions, but not the published version with copy-editing, proofreading, and formatting added by IEEE. For more information, please go to: http://www.ieee.org/publications_standards/publications/rights/paperversionpolicy.html. Permission to reprint/republish this material for commercial, advertising, or promotional purposes or for creating new collective works for resale or redistribution must be obtained from IEEE by writing to the IEEE Intellectual Property Rights Office, 445 Hoes Lane, Piscataway, NJ 08854-4141 or pubs-permissions@ieee.org. Copyright © 2019 IEEE. All rights reserved.

Abstracting and Library Use: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy for private use of patrons, provided the percopy fee indicated in the code at the bottom of the first page is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Unsubscribe: If you no longer wish to receive this ComputingEdge mailing, please email IEEE Computer Society Customer Service at help@computer.org and type "unsubscribe ComputingEdge" in your subject line.

IEEE prohibits discrimination, harassment, and bullying. For more information, visit www.ieee.org/web/aboutus/whatis/policies/p9-26.html.

IEEE Computer Society Magazine Editors in Chief

Computer David Alan Grier (Interim), Djaghe LLC

IEEE Software Ipek Ozkaya, Software Engineering Institute

IEEE Internet Computing George Pallis, University of Cyprus

IT Professional Irena Bojanova, *NIST* *IEEE Security & Privacy* David Nicol, University of Illinois at Urbana-Champaign

IEEE Micro Lizy Kurian John, University of Texas, Austin

IEEE Computer Graphics and Applications Torsten Möller, University of Vienna

IEEE Pervasive Computing Marc Langheinrich, *University of Lugano* Computing in Science & Engineering Jim X. Chen, George Mason University

IEEE Intelligent Systems V.S. Subrahmanian, Dartmouth College

IEEE MultiMedia Shu-Ching Chen, Florida International University

IEEE Annals of the History of Computing Gerardo Con Diaz, University of California, Davis



How Do We Prepare the Next Generation for a Career in Our Digital Era? Service Learning in Engineering Education

Project-Based Learning Inspires Cybersecurity Students: A Scholarship-for-Service Research Study



40 Thoughts on Cyberbullying

Careers

- 8 How Do We Prepare the Next Generation for a Career in Our Digital Era?
- 11 Service Learning in Engineering Education NICHOLAS J. KIRSCH

Artificial Intelligence

- 17 Are Robots Taking Our Jobs? A RoboPlatform at a Bank PRZEMYSLAW LEWICKI, JACEK TOCHOWICZ, AND JEROEN VAN GENUCHTEN
- 21 A Budding Romance: Finance and AI XIAO-PING STEVEN ZHANG AND DAVID KEDMEY

Internet

26 Project-Based Learning Inspires Cybersecurity Students: A Scholarship-for-Service Research Study

ALAN T. SHERMAN, PETER A.H. PETERSON, ENIS GOLASZEWSKI, EDWARD LAFEMINA, ETHAN GOLDSCHEN, MOHAMMED KHAN, LAUREN MUNDY, MYKAH RATHER, BRYAN SOLIS, WUBNYONGA TETE, EDWIN VALDEZ, BRIAN WEBER, DAMIAN DOYLE, CASEY O'BRIEN, LINDA OLIVA, JOSEPH ROUNDY, AND JACK SUES

34 Learning to Network STEPHEN D. CROCKER

Social Media

- 40 Supervised Learning for Fake News Detection JULIO C. S. REIS, ANDRE CORREIA, FABRICIO MURAI, ADRIANO VELOSO, AND FABRICIO BENEVENUTO
- 46 Thoughts on Cyberbullying NIR KSHETRI AND JEFFREY VOAS

Departments

- 4 Magazine Roundup
- 7 Editor's Note: Engineering Your Career
- 72 Conference Calendar

Subscribe to *ComputingEdge* for free at **www.computer.org/computingedge.**

Magazine Roundup

he IEEE Computer Society's lineup of 12 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 visualization and microchip design. Here are highlights from recent issues.

Computer

The Harmonization of Accessibility Standards for Public Policies

Today, individuals can access an ever-increasing number of services via the Internet. However,

October 2019

only when all people are able to completely access the Internet can a digital society be considered universal. The authors of this article from the July 2019 issue of *Computer* present a proposal to harmonize accessibility standards that all countries must adhere to.

Computing in Science & Engineering

Lifecycle Support for Scientific Investigations: Integrating Data, Computing, and Workflows

Scientific workflows have emerged as a model for representing the complex processes carried out by scientists

throughout their investigations, encompassing research activities corresponding to data collection, data flow, computation, output analysis, and all the ways these are used together to produce results. Existing infrastructures support elements of the workflow, such as data repositories or computing services, but these are not integrated as interactive environments that provide full investigation lifecycle support. The Digital Environment for Enabling Data-driven (DEEDS) Sciences project brought together domain scientists and computer scientists to create a platform that provides interactive end-to-end support for diverse scientific workflows. Key among requirements were preservation, provenance, coupling of data and computing, results traceability, collaborative sharing, exploration, and publication of the full products of research work. This article from the July/August 2019 issue of Computing in Science & Engineering highlights use cases that

contributed to DEEDS development and concludes with lessons learned from a process that joined experiences and perspectives from diverse science domains.

IEEE Annals of the History of Computing

Polish Text Editors during the Fall of the Iron Curtain

The Polish economic transformation of the 1990s created an appetite for software that was only partially satisfied by piracy. This market was yet to be taken seriously by Western companies, so local developers stepped up to fill the void. They translated obscure foreign applications, created weird character encoding standards, and built complex business software from scratch, shaping the local IT market for years to come. Read more in the April–June 2019 issue of IEEE Annals of the History of Computing.

IEEE Computer Graphics and Applications

Personalized Sketch-Based Brushing in Scatterplots

Brushing is at the heart of most modern visual analytics solutions, and effective and efficient brushing is crucial for successful interactive data exploration and analysis. As the user plays a central role in brushing, several datadriven brushing tools have been designed that are based on predicting the user's brushing goal. All of these general brushing models learn the user's average brushing preference, which is not optimal for every user. In this article from the July/August 2019 issue of IEEE Computer Graphics and Applications, the authors propose an innovative framework that offers the user opportunities to improve the brushing technique while using it. They realized this framework with a CNN-based brushing technique, and the result shows that with additional data from a particular user, the model can be refined (better performance in terms of accuracy), eventually converging to a personalized model based on a moderate amount of retraining.

IEEE Intelligent Systems

Autonomous Intelligent Agents for Team Training: Making the Case for Synthetic Teammates

The rise in autonomous system research and development combined with the maturation of computational cognitive architectures holds the promise of high-cognitive-fidelity agents capable of operating as team members for training. In this article from the March/April 2019 issue of IEEE Intelligent Systems, the authors report an ACT-R model capable of operating as a team member within a remotely piloted aerial system. They provide results from a first-of-its-kind controlled, randomized empirical evaluation in which teams that worked with an AST were compared against all-human teams. The results demonstrate that ASTs can be incorporated into human teams, providing training opportunities when teammates are unavailable. The authors conclude with issues faced in developing ASTs and lessons learned for future and current developers.

IEEE Internet Computing

Seamless Virtualized Controller Migration for Drone Applications

The authors of this article from the March/April 2019 issue of *IEEE Internet Computing* consider a virtualized edge-computing infrastructure for drone applications, in which a virtualized container running on an edge node controls drones and a software-defined network provides a network connectivity between the drones and their virtualized controllers. The authors propose a seamless migration scheme that migrates a virtualized drone controller to an edge node that is close to its associated drone without suspending the drone control in the edge-computing infrastructure.

IEEE Micro

Composable Building Blocks to Open Up Processor Design

In this article from the May/June 2019 issue of *IEEE Micro*, the authors present a framework called Composable Modular Design (CMD) to facilitate the design of out-of-order processors. In CMD, 1) the interface methods of modules provide instantaneous access and perform atomic updates to the state elements inside the module; 2) every interface method is guarded, i.e., it cannot be applied unless it is ready; and 3) modules are composed by atomic rules that

call interface methods of different modules. A rule either successfully updates the state of all the called modules or does nothing. The atomicity properties of interfaces in CMD ensure composability when modules are refined selectively. The authors show the efficacy of CMD by building an outof-order RISC-V processor, which boots Linux. Modules designed using CMD (e.g., ROB, load-store unit, etc.) can be used and refined by other implementations.

IEEE MultiMedia

ToothPic: Camera-Based Image Retrieval on Large Scales

Being able to reliably link a picture to the device that shot it is of paramount importance to give credit or assign responsibility to the author of the picture itself. However, this task needs to be performed at large scales due to the recent explosion in the number of photos taken and shared. Existing methods cannot satisfy those requirements. Methods based on the photo response nonuniformity (PRNU) of digital sensors are able to link a photo to the device that shot it and have already been used as proof in the court of law. Such methods are reliable but so far, they can only be used for small-scale forensic tasks involving few cameras and pictures. ToothPic, an acronym for "Who Took This Picture?," is a novel image retrieval engine that allows for finding all the pictures in a large-scale database shot by a given query camera. Read more in the April–June 2019 issue of *IEEE MultiMedia*.

IEEE Pervasive Computing

A User Study of Semi-Autonomous and Autonomous Highway Driving: An Interactive Simulation Study

The aim of the study presented in this article from the January-March 2019 issue of IEEE Pervasive Com*puting* is to explore user acceptance of semi-autonomous and fully autonomous vehicles on a highway through the use of an interactive simulator. Participants were asked to experience driving modes with three levels of autonomy and complete questionnaires with items selected from traditional and automotive-specific technology acceptance models. The three levels of automation were manual driving (no automation as a baseline condition), semi-autonomous driving where drivers were able to indicate lane-change decisions, and fully autonomous driving. Results indicate that, within the limited experience of the interactive simulation, users grew to like the automated system as much as manual control during later portions of the study. Overall, this work suggests that the driver will quickly grow to like automated driving features and may rapidly become less anxious about the loss of control experienced.

IEEE Security & Privacy

A Cognitive Protection System for the Internet of Things

Conventional cybersecurity neglects

the Internet of Things' physicality. The authors of this article from the May/June 2019 issue of *IEEE Security & Privacy* propose a cognitive protection system capable of using system models to ensure command safety while monitoring system performance. They develop and test a cognitive firewall and cognitive supervisor. This system is tested in theory and practice for three threat models.

IEEE Software

User Engagement in the Era of Hybrid Agile Methodology

Contemporary software development and implementation projects are increasingly adopting agile methods by tailoring and blending agile techniques into a traditional project framework. Common tailoring methods employed by project teams emphasize flexibility to embrace local project context. Read more in the July/August 2019 issue of *IEEE Software*.

IT Professional

Exploiting Edge Computing for Privacy-Aware Tourism Demand Forecasting

Taking advantage of the processing power that modern smartphones possess, the authors of this article from the May/June 2019 issue of *IT Professional* advocate for a privacy-aware approach to predict and suggest places of interest for travelers. This approach solves a significant privacy flaw that exists in prevalent tourism applications. •

Engineering Your Career

ccording to the US Bureau of Labor Statistics, employment in the fields of computing and information technology is expected to grow faster than most other occupations, with thousands of new jobs created every year. To meet such large workforce demands, we need to encourage young people to build a career in the field and help set them up for success in the workplace. This issue of *ComputingEdge* focuses on accomplishments and challenges in computing education and job training.

In *Computer*'s "How Do We Prepare the Next Generation for a Career in Our Digital Era?," the author asserts that soft skills such as adaptability and creativity are just as important as technical skills in modern work environments. Meanwhile, *IEEE Pervasive Computing*'s "Service Learning in Engineering Education" describes a teaching method that aims to motivate students to choose computer-engineering careers.

Many people wonder how artificial intelligence (AI) will affect the careers of the future. *IEEE Software*'s "Are Robots Taking Our Jobs? A RoboPlatform at a Bank" describes how banking software performs administrative work faster and more accurately than employees do. "A Budding Romance: Finance and AI," from *IEEE MultiMedia*, covers the ways in which the financial industry is using AI to create operational efficiencies and make market predictions.

Knowledge of the Internet is necessary in many of today's occupations. *IEEE Security & Privacy*'s "Project-Based Learning Inspires Cybersecurity Students: A Scholarship-for-Service Research Study" details a student project that exposed security vulnerabilities in a university's network. The author of "Learning to Network," from *IEEE Annals of the History of Computing*, recounts his role in developing the Internet.

Finally, this *ComputingEdge* issue includes two articles on social media. *IEEE Intelligent Systems*' "Supervised Learning for Fake News Detection" discusses approaches for identifying fake news stories that are disseminated on social media. *Computer*'s "Thoughts on Cyberbullying" pushes for increased efforts to address online harassment from regulators, law enforcement, educators, and tech companies. \bigcirc

COMPUTING EDUCATION



How Do We Prepare the Next Generation for a Career in Our Digital Era?

Louise M. Morman, Lockheed Martin Leadership Institute

Technical skills alone won't be enough for career success in the digital age. It's much more than teaching people to code. Instead, success will hinge on critical soft skills and the RISC– resiliency, inner strength, strategic thinking, and a collaborative spirit—digital mind–set abilities.

transformation" of organizations. Digital transformation is not fragmented digitization; it is a complete rethinking of the overall business model with a customer-driven emphasis supported by the use of digital technologies throughout the business process.

Are businesses prepared for this transition? In a 2016 MIT Sloan Management Review/Deloitte University Press study¹ of more than 3,700 executives, managers, and analysts around the globe, 90% of respondents recognized that digi-

hether it is at the World Economic Forum or in the suites of CEOs, there is strong agreement that artificial intelligence, machine learning, mobile technologies, robotics, biological technologies, quantum computing, and other technological innovations require the "digital

Digital Object Identifier 10.1109/MC.2019.2903328 Date of publication: 14 May 2019 tal trends would moderately or greatly disrupt their industries. Only 44%, however, believed that their organizations were prepared for the disruption.

Another insight from the MIT/Deloitte report is that the companies that appear more successful in their digital transformation efforts put their focus on soft skills. In response to a question about the most important skill for leaders to succeed in a digital environment, only 18% of respondents listed technological skills as most

8

EDITOR ANN E.K. SOBEL Miami University; sobelae@muohio.edu

important. Instead, they highlighted attributes such as having a transformative vision (22%), being a forward thinker (20%), having a change-oriented mind-set (18%), or other leadership and collaborative skills (22%). Employers are realizing more and more that career success hinges on critical soft skills—the things that computers don't do as well as humans.

Changes in what Klaus Schwab, executive chair of the World Economic Forum, calls the Fourth Industrial Revolution will impact a very broad range of the workforce. Work in the future will be altered in ways we haven't seen in the past, and changes will reach well beyond hourly workers. All kinds of occupations will see change, including engineers, accountants, coders, and surgeons. Computers are very effective at handling logical and process-oriented activities. The work of knowledge workers and college graduates will be changed, so people will have to constantly reinvent themselves and deal with a never-ending change journey throughout their careers.

It is often said that difficulties implementing digital transformation have been more about people dealing with change than with the technological tools. And I believe this may be the most significant issue of all. We will need to find answers to the following questions. What kinds of mind-sets do people need to thrive in the digital era? How do they need to *be*? And what are we doing to prepare now?

These answers will incorporate abilities in what I call the RISC digital mindset, which focuses on Resiliency, Inner strength, Strategic thinking, and Collaborative spirit (see "RISC Digital Mindset").

Many people talk about culture change as something that happens only in big companies. But welcoming change is a talent that takes time and effort for an individual to develop. The RISC mindset is very personal and a lifelong development effort. My belief is that we should start that transformational leadership development in undergraduate postsecondary education.

The fact that millennials and Generation Z grew up in a digital environment does not mean that they embrace change and are equipped to navigate in an uncertain work world. They have been conditioned to expect clear direction, which has shielded them from the messiness of the real world. Universities must be accountable for the resiliency of the students they graduate, now more than ever. Technical skills alone cannot guarantee a successful employee in the digital age.

What can be done to address obtaining these abilities in colleges? One way to incorporate transformational

RISC DIGITAL MINDSET

Sesiliency-Adaptability

- » comfort in uncertainty and unstructured environments
- » navigating through complexity, volatility, and ambiguity
- » embracing change, especially when the change isn't our idea
- » curiosity and growth mind-set
- » reinventing ourselves as a way of life
- » learning from failure.

nner strength-Being your best evolved self

- » self-awareness
- » open mindedness and inclusion
- » courage
- » humility-keeps us in touch with all we don't know
- » embodying our humanity
- » overcoming fear of conflict and risk.

Ctrategic thinker-Holistic



- » creative thinking
- » big picture and systems thinking
- » customer obsessed
- » forward thinker with transformative vision
- » ideas based on intuition and insights as well as data and history
- » openness to new ideas and possibilities.

Ollaborative spirit

- collaborative problem solving
- » emotional intelligence
- » heart.

COMPUTING EDUCATION

learning is by partnering with leadership development professionals from industry and executive coaches. Transformation is intensely personal, takes time, and requires coaching from people with the proper training and experience. More than seven years ago, we created the Lockheed Martin Leadership Institute in the College of Engineering and Computing at Miami University to do just that. Ours is an example of an intensive three-year program for a targeted group of students. And there are other universities that have created excellent large-scale programs, such as the Doerr Institute for New Leaders at Rice University.

It is also important for the university to invest in transformational leadership development for faculty and staff. If faculty members have engaged in their own transformational process, they will be better able to understand the changes the students experience and can remain relevant in the future.

y hope is that universities will make this a priority and start preparing the next generation today for the resiliency and adaptability, inner strength, strategic thinking, and collaborative spirit that are essential for the coming digital era.

REFERENCE

 G.C. Kane, D. Palmer, A. N. Phillips, D. Kiron, and N. Buckley, "Aligning the organization for its digital future," *MIT Sloan Management Rev.*, July 2016. [Online]. Available: https:// sloanreview.mit.edu/digital2016

LOUISE M. MORMAN is the executive director of the Lockheed Martin Leadership Institute. Contact her at mormanIm@miamioh.edu.

This article originally appeared in Computer, *vol. 52, no. 5, 2019.*

stay on the Cutting Edge of Artificial Intelligence



IEEE Intelligent Systems provides peerreviewed, cutting-edge articles on the theory and applications of systems that perceive, reason, learn, and act intelligently.

> Intelligent Systems



IEEE MultiMedia serves the community of scholars, developers, practitioners, and students who are interested in multiple media types and work in fields such as image and video processing, audio analysis, text retrieval, and data fusion.

Read It Today! www.computer.org/multimedia

Digital Object Identifier 10.1109/MC.2019.2911840

The #1 Al Magazine

Digital Object Identifier 10.1109/MC.2019.2911839

Service Learning in Engineering Education

Nicholas J. Kirsch University of New Hampshire

Department editor: Andrew L. Kun; andrew.kun@unh.edu EPICS in IEEE empowers students to work with local service organizations to apply technical knowledge to implement solutions for a community's unique challenges. In this way, EPICS in IEEE not only assists communities in achieving their specific local improvement goals but also encourages students to

pursue engineering for community improvement as a career.

There are efforts throughout the world to broaden participation in science, technology, engineering, and mathematics (STEM) to meet increasing workforce demands. Possibly an even greater benefit of these efforts is that a diverse workforce will be able to propose a diverse set of ideas to solve future problems.¹ However, broadening participation in engineering is a challenge because people associate many negative stereotypes with engineering, which discourages some students from choosing the field. Further, engineering is a challenging topic to study. As such, there are many programs and pedagogies in place to overcome these challenges.²⁻⁴ Service learning is one approach that has tremendous potential to broaden participation, break down stereotypes, and increase student performance.

Service learning is an experiential-based teaching method that couples academic work with community service projects. Students go out in their community to solve a real-world problem.⁵ The problem becomes the basis for a project, which in turn provides hands-on experiences for students to hone theoretic and practical skills. At the same time, the community partners receive a solution to a problem. Service learning projects are opportunities for engineers to solve problems outside of traditional topics and, similarly, demonstrate to communities that engineers do not just do stereotypical work.

EPICS IN IEEE

While many academic programs have implemented service learning, Purdue University in 1995 created Engineering Projects in Community Service (https://engineering.purdue.edu/EPICS), which includes curriculum and service learning best practices for engineering education.⁶ In EPICS projects, students work on engineering-related, interdisciplinary projects with local non-profit organizations (NPOs). There is typically a multidisciplinary approach to broadly solve problems and vertically integrate learners; students with a range of knowledge and abilities learn

from one another. This approach benefits students by giving them a more diverse group of people to teach and learn from. Finally, EPICS projects aim to have start-to-finish design: they do not terminate at a proof-of-concept but rather can be multiyear efforts that go through many iterations and design cycles.

In 2009, IEEE leveraged its more than 400,000 members to expand this service learning education model throughout the world by starting EPICS in IEEE. This program encourages student branches to partner with local NPOs to provide a solution to a challenge within their community. Projects can be proposed by a student branch but are typically advised by a university professor or another IEEE member. The student volunteers then work directly with the NPO and their advisors to use their engineering skills to solve the problem. EPICS in IEEE has a unique approach in encouraging the participation of K–12 grade students to work alongside university students (see Figure 1). This approach, known as vertical integration, has been shown to lead to positive outcomes for participants of all ages.⁷ Younger learners get exposed to real-world applications of engineering and basic technical skills, while the university participants can refine their skills by teaching the younger learners.



Figure 1. University of New Hampshire students and local high school volunteers test a water flow sensor for "A Wireless Sensor Network to Restore Oysters in the Great Bay of New Hampshire." This EPICS in IEEE project exemplifies the vertical integration of learners.

Over the last eight years, EPICS in IEEE has provided more than \$500,000 to fund 96 projects in 34 countries. More than 200,000 people have been impacted by this program, either directly as volunteers or indirectly as community members. Further, more than 1,500 of the pre-university students are women.

Typically, EPICS in IEEE provides about \$5,000 per project to cover the cost of materials, supplies, and equipment, though some projects receive more than \$20,000. The program funds a wide variety of projects on a rolling basis year-round, and generally the projects fall into one of four categories:

- Access and Abilities—Access and Abilities projects help enable adaptive services, clinics for those in need (such as children with disabilities), programs for adults, and assistive technologies.
- *Education and Outreach*—EPICS in IEEE strives to help young students discover the benefits of STEM for their futures. Many projects give students hands-on experiences to stimulate interest in those fields. Through these projects, communities and schools lacking strong engineering programs gain new curriculums along with new facilities to explore new areas of a topic.

- *Environment*—Engineering and science are key to meeting environmental challenges. In communities around the world, environments change with the evolution of technology and the need for sustainability. Many EPICS in IEEE projects concern themselves with recycling, as well as with new ways to create electricity and energy, including the use of renewable energy sources. Through these projects, young students learn about the impact of environmental issues and how engineering can help resolve them. They also gain exposure to potential jobs given the growing demand for alternative energy and environmental solutions.
- *Human Services*—Through their experiences in these EPICS in IEEE projects, students find connections between engineering and the tremendous scope of community needs globally. This includes homelessness prevention, affordable housing, family and children agencies, neighborhood revitalization, and local government. Even after Human Services project is complete, lasting impact continues to be felt through the local non-profit organization's involvement.

Administratively, running EPICS in IEEE can be challenging given IEEE's global reach. The diverse collection of countries and education systems involved in the program requires flexible administration. Seemingly "simple" tasks such as project reporting become more complex when considering the differences in academic year calendars throughout the world. However, due to IEEE's established framework, the various financial and bureaucratic mechanisms for research funding projects are streamlined by directly funding the parent IEEE Section in which the project is administered.

The following is a sample of EPICS in IEEE projects that have had an impact in different regions of the world.

Virtual Reality Vision Therapy

Binocular dysfunction, the inability of both eyes to work together, has been called the hidden learning disability. This condition drastically affects children's academic performance and rivals other disabling physical conditions, including brain injury. Treatment is expensive, not typically covered by insurance, and requires visiting a therapist. A team of four biomedical engineering students from the New Jersey Institute of Technology (NJIT) and two high school students collaborated with The Eye Institute, a nonprofit organization, to create a significantly less expensive solution: a virtual reality (VR), home-based device designed to make therapy sessions both effective and fun.



Figure 2. An image of the VR system developed by students at NJIT to rehabilitate children with binocular dysfunction.

With the help of an EPICS in IEEE grant, the group transformed at-home vision therapy into a high-quality, 3D video game that appeals to children. Patients undergo therapy sessions wearing a head-mounted display that provides a colorful, engrossing VR experience. By correctly aligning and maintaining eye position for an amount of time set by the clinician, the patient "destroys" 3D digital models of alien creatures (see Figure 2).

As they take part in the overall design process, the student volunteers on this project gain practical, hands-on experiences that expand their knowledge of engineering principles. Once completed and validated, the game will cost about \$600—a much more affordable option for families with limited financial resources. Ultimately, the students want to make it available to all children who need vision therapy.

Enabling "Casual Talk" with the Deaf

In India, the creative blending of three technologies provided the framework for empowering people with hearing disabilities to have "casual" or simple conversations with those who can hear. The IEEE Student Branch at National Engineering College in Kovilpatti, Tamil Nadu, used an EPICS in IEEE grant of \$7,900 to develop "CasTalk." The project relied on smart mobile devices with 3G or 4G technologies, an animated video streamer, and cloud resources. The eventual result will be simple, natural communication between the hearing disabled and people with normal hearing who live in South India where the regional sign language is a combination of English and Tamil (see Figures 3 and 4).



Figure 3. One of the IEEE student volunteers developing the CasTalk system.



Figure 4. A screenshot from CasTalk analyzing different attributes of a video feed.

Controlling Insect Infestations with Sensor Networks

While talking with local farmers about insects that were destroying their crops, IEEE students at the University of Johannesburg in South Africa had an idea: they wanted to develop a method to help farmers detect and track the movements of insects to help increase the effectiveness of pesticides.

Partnering with the South African Subtropical Growers Association, which represents farmers who grow avocados, mangos, and litchi and macadamia nuts, the students designed and built a harmonic radar system that tracks small insects by attaching an RFID tag to a captured bug and releasing it. The radar can determine the RFID tag's location even in a cluttered environment. The tag uses the original radar signal as an energy source, reemitting a harmonic of the transmitting signal. Tuning the receiver to the harmonic frequency distinguishes the tagged target—and the others in its cluster—from background clutter. The students analyzed signal propagation and harmonic effects of the system to increase the range. This project directly benefited the nonprofit partner by creating a solution to its problem.

CONCLUSION

EPICS in IEEE is having a unique impact across the globe, not only helping educate young learners and university students through service learning but providing technological solutions for communities and NPOs with varying needs. The program truly exemplifies IEEE's motto: "Advancing Technology for Humanity." We encourage participation in our program to improve learning and to solve community problems. For more information on EPICS in IEEE, visit http://epics.ieee.org.

REFERENCES

- N. Dasgupta and J.G. Stout, "Girls and Women in Science, Technology, Engineering, and Mathematics: STEMing the Tide and Broadening Participation in STEM Careers," *Policy Insights from the Behavioral and Brain Sciences*, vol. 1, no. 1, 2014, pp. 21–29.
- 2. N.L. Fortenberry and J.J. Powlik, "Helping to Shape the Future of Education," *IEEE Trans. Education*, vol. 40, no. 4, 1997; doi.org/10.1109/13.759675.
- K.W. Jablokow, "Engineers as Problem-Solving Leaders: Embracing the Humanities," IEEE Technology and Society Mag., vol. 26, no. 4, 2007, pp. 29–35.
- 4. P.M. Jansson et al., "Creating an Agile ECE Learning Environment through Engineering Clinics," *IEEE Trans. Education*, vol. 53, no. 3, 2010, pp. 455–462.
- R.G. Bringle and J.A. Hatcher, "Implementing Service Learning in Higher Education," *J. Higher Education*, vol. 67, no. 2, 1996, pp. 221–239.
- W.C. Oakes, E.J. Coyle, and L.H. Jamieson, "EPICS: A Model of Service-Learning in an Engineering Curriculum," *Proc. 2000 ASEE Ann. Conf. and Exposition*, 2000, pp. 2623–2636.
- 7. F. Giralt et al., "Two Way Integration of Engineering Education through a Design Project," *J. Engineering Education*, vol. 89, no. 2, 2000, pp. 219–229.

ABOUT THE AUTHOR

Nicholas J. Kirsch is an associate professor in the Department of Electrical and Computer Engineering at the University of New Hampshire. Contact him at nicholas.kirsch@unh.edu.

This article originally appeared in IEEE Pervasive Computing, *vol. 17, no. 2, 2018.*



PURPOSE: The IEEE Computer Society is the world's largest association of computing professionals and is the leading provider of technical information in the field.

MEMBERSHIP: Members receive the monthly magazine *Computer*, discounts, and opportunities to serve (all activities are led by volunteer members). Membership is open to all IEEE members, affiliate society members, and others interested in the computer field.

COMPUTER SOCIETY WEBSITE: www.computer.org

OMBUDSMAN: Direct unresolved complaints to ombudsman@computer.org.

CHAPTERS: Regular and student chapters worldwide provide the opportunity to interact with colleagues, hear technical experts, and serve the local professional community.

AVAILABLE INFORMATION: To check membership status, report an address change, or obtain more information on any of the following, email Customer Service at help@computer.org or call +1 714 821 8380 (international) or our toll-free number,

- +1 800 272 6657 (US):
 - Membership applications
 - Publications catalog
 - Draft standards and order forms
 - · Technical committee list
 - Technical committee application
 - Chapter start-up procedures
 - Student scholarship information
 - Volunteer leaders/staff directory
 - IEEE senior member grade application (requires 10 years practice and significant performance in five of those 10)

PUBLICATIONS AND ACTIVITIES

Computer: The flagship publication of the IEEE Computer Society, *Computer* publishes peer-reviewed technical content that covers all aspects of computer science, computer engineering, technology, and applications.

Periodicals: The society publishes 12 magazines, 15 transactions, and two letters. Refer to membership application or request information as noted above.

Conference Proceedings & Books: Conference Publishing Services publishes more than 275 titles every year.

Standards Working Groups: More than 150 groups produce IEEE standards used throughout the world.

Technical Committees: TCs provide professional interaction in more than 30 technical areas and directly influence computer engineering conferences and publications.

Conferences/Education: The society holds about 200 conferences each year and sponsors many educational activities, including computing science accreditation.

Certifications: The society offers three software developer credentials. For more information, visit www.computer.org/certification.

2019 BOARD OF GOVERNORS MEETING

25 - 27 October: Hilton McLean Tysons Corner, McLean, VA

EXECUTIVE COMMITTEE

President: Cecilia Metra

President-Elect: Leila De Floriani Past President: Hironori Kasahara First VP: Forrest Shull; Second VP: Avi Mendelson; Secretary: David Lomet; Treasurer: Dimitrios Serpanos; VP, Member & Geographic Activities: Yervant Zorian; VP, Professional & Educational Activities: Kunio Uchiyama; VP, Publications: Fabrizio Lombardi; VP, Standards Activities: Riccardo Mariani; VP, Technical & Conference Activities: William D. Gropp

2018–2019 IEEE Division V Director: John W. Walz 2019 IEEE Division V Director Elect: Thomas M. Conte 2019–2020 IEEE Division VIII Director: Elizabeth L. Burd

BOARD OF GOVERNORS

Term Expiring 2019: Saurabh Bagchi, Gregory T. Byrd, David S. Ebert, Jill I. Gostin, William Gropp, Sumi Helal

Term Expiring 2020: Andy T. Chen, John D. Johnson, Sy-Yen Kuo, David Lomet, Dimitrios Serpanos, Hayato Yamana

Term Expiring 2021: M. Brian Blake, Fred Douglis, Carlos E. Jimenez-Gomez, Ramalatha Marimuthu, Erik Jan Marinissen, Kunio Uchiyama

EXECUTIVE STAFF

Executive Director: Melissa A. Russell Director, Governance & Associate Executive Director: Anne Marie Kelly Director, Finance & Accounting: Sunny Hwang Director, Information Technology & Services: Sumit Kacker Director, Marketing & Sales: Michelle Tubb Director, Membership Development: Eric Berkowitz

COMPUTER SOCIETY OFFICES

Washington, D.C.: 2001 L St., Ste. 700, Washington, D.C. 20036-4928; Phone: +1 202 371 0101; Fax: +1 202 728 9614; Email: help@computer.org Los Alamitos: 10662 Los Vaqueros Cir., Los Alamitos, CA 90720; Phone: +1 714 821 8380; Email: help@computer.org

MEMBERSHIP & PUBLICATION ORDERS

Phone: +1 800 678 4333; Fax: +1 714 821 4641; Email: help@computer.org

IEEE BOARD OF DIRECTORS

President & CEO: Jose M.D. Moura President-Elect: Toshio Fukuda Past President: James A. Jefferies Secretary: Kathleen Kramer Treasurer: Joseph V. Lillie Director & President, IEEE-USA: Thomas M. Coughlin Director & President, Standards Association: Robert S. Fish Director & VP, Educational Activities: Witold M. Kinsner Director & VP, Membership and Geographic Activities: Francis B. Grosz, Jr.

Director & VP, Publication Services & Products: Hulya Kirkici Director & VP, Technical Activities: K.J. Ray Liu





Editor: Michiel van Genuchten VitalHealth Software genuchten@ieee.org



Editor: Les Hatton Oakwood Computing Associates lesh@oakcomp.co.uk

Are Robots Taking Our Jobs? A RoboPlatform at a Bank

Przemysław Lewicki, Jacek Tochowicz, and Jeroen van Genuchten

From the Editors

Robots dominate some Hollywood movies and economic journals. This article describes the impact of software robots in a Polish bank. —*Michiel van Genuchten and Les Hatton*

AUTOMATION HAS BEEN replacing manual activities in workplace for decades. Robots had been most active in blue-collar industrial manufacturing. Now that they are entering white-collar jobs, in the form of software taking over administrative work, it is time to ask the following questions.

- Are these real robots?
- Why do we need them?
- Are they going to take over all jobs once self-learning algorithms mature?

We will try to answer these questions by analyzing the deployment of the self-build RoboPlatform at ING Slaski Bank in Poland.

Robotic Process Automation

Robotic process automation (RPA) mimics human administrative actions.

The software robot receives a digital form or email and processes the requests by following a script. It reads the incoming data, opens screens, and enters data, just as a human would. The robots can operate in the following two modes.

- *Attended*: A robotic digital assistant resides on the desktop, and the human employee can trigger it to work on repetitive, mundane tasks while he or she works on other things.
- *Unattended*: The robot works autonomously under its own credentials on scheduled tasks.

RPA is a booming business. Everest Group² states that RPA adoption exceeded 100% growth in 2017, buoyed by new buyers of all sizes and industries. Forrester³ predicted that the RPA market, which was only US\$250 million in 2016, would grow to US\$2.9 billion in 2021. As is the case for other companies, ING wants to increase the speed and accuracy of its processes, enhance the customer experience, and reduce costs. In business processes, system limitations are difficult to overcome and stretch across many applications in the organization. ING is investing in global core banking platforms, but these complex transition programs take years to accomplish. In the meantime, you can deliver digital solutions with RPA, and leave your legacy software and business processes largely untouched.

Use Cases

ING Slaski, a bank in Poland that has 4.5 million customers, started to use RPA 10 years ago. Back then, the need for an end-user computing platform led to the implementation of Macro-Platform. From then on, staff at Slaski could outsource the task of retyping the details for a personal loan or current account from the mainframe

Digital Object Identifier 10.1109/MS.2019.2897337 Date of publication: 16 April 2019

application into an Excel file to a MacroPlatform script. With the successor RoboPlatform, the customer support specialist can have an engaging conversation with the customer on the phone while the robot retrieves all the customer data from various applications and presents it in one overview. RoboPlatform increases efficiency and reaction speed; on average, it completes scripts 5.5 times faster than an employee would. With more than 1,600 scripts in production, the impact for Slaski's operations is considerable. More than 700 of Slaski's 1,100 total operations and customer support employees use RoboPlatform as part of their daily activities. Since the robotic capacity is the same as 70 fulltime equivalents' manual work, the bank was able to grow its business and comply with increasing regulatory demands without hiring additional staff. In addition to the improved speed and efficiency, the robots are more accurate also, assuming the virtual machines are running and the scripts were developed correctly. Robots work on customer due diligence tasks dutifully and can log and store all required auditable data without mistakes. The robots do not have bad days, let alone hangovers or broken hearts. RoboPlatform also is an integration solution between applications. Output products from applications can be passed on to the business lending backoffice application using RoboPlatform.

Make or Buy?

There were no standard RPA products on the market 10 years ago. The initial version used Pascal as a programming language for the product and the scripts. Two years ago, MacroPlatform was fully rebuilt using .NET technology and transferred into the RoboPlatform RPA product. Scripting can now be done in Pascal, C#, or Visual Basic, and the RoboPlatform is 140 KLOC.

When Slaski decided to rebuild MacroPlatform into RoboPlatform, we took a close look at the solutions that were available to purchase. The biggest disadvantage of moving to an external vendor was that there was no way to incorporate the existing MacroPlatform scripts into these external products without rebuilding them almost completely. Since our business is banking, the highest standards for security and credential management are necessary. The new product should be enterprise ready, such as supporting role-based access by integrating with our central directory services (Microsoft Active Directory). We concluded that the external RPA vendors were not mature enough at the time. Also, the yearly license fees for the large number of robots needed could be more than US\$1 million, which would destroy our return on investment. Therefore, we decided to leverage on our RPA experience.

We built most of the RoboPlatform components inhouse, including state machine implementation and workflow implementation. We built a full debugger as well as a simple code review tool. We developed components for multisession mode, for when the scripts do not require a graphical interface. There is risk assessment and monitoring of scripts to identify errors and irregularities, and a screenshot capturing tool is active during the session.

We use about 200 open libraries. The most important are FreeRDP, AvalonDock, CefGlue, MaterialDesign + MahApps.Metro, Selenium, Microsoft.Build, Rolsyn, NRefactory, Fusion, Microsoft IUIAutomation, Reactive Extensions, FluentScheduler, Vault, EntityFramework, Dapper, JQuery, Bootstrap, Bootstrap, SinglarR, AutoMapper, Newtonsoft. Json, Caliburn.Micro, SimpleInjector, LightInject, and Accord.

RoboPlatform consists of three main modules.

- MachineHeartbeat: This Windows service is responsible for managing the machines on which the robots are running. It keeps track of the machines' activity status and monitors queued tasks that need to be processed in unattended mode, to make sure that the machines are ready and available. It manages logging into the Robot account, and the password that is downloaded from the password vault.
- *Engine:* The console application is responsible for the correct execution of C#, VB, Pascal, and Workflow scripts. The engine also sends logs to BotSlave.
- *BotSlave*: This console application starts Bots (a script or set of scripts) after a successful login. Its task is to start the Engine application in a timely manner, to log the actions performed by running the script, and to log out of the robot account after the task has been completed.

Robot Resource System

Our unattended robots do not get a salary, access badge, or holidays, but they need authorization to be able to work in specific applications, just like humans. They also need a manager who is accountable for their actions. Therefore, we introduced the Robot Resource System (RRS), built in .NET, where all robot accounts that are in use within ING globally are registered (Figure 1). The RRS feeds all of the applications that are already in place for identity access management. The RRS connects to a password vault where the robots' passwords are securely stored.

IMPACT

RoboPlatform will fetch the passwords from the vault when the script requires a login into the network. In this way, no human needs to know the password for the robot account, unless there is an emergency (e.g., a robot malfunctioning).

Agile

Multiple roles within the organization use the RoboPlatform ecosystem. The end user triggers scripts, the IT custodian and operations manager monitor performance by using the dashboard, and the scheduler plans the work of unattended robots. The process owner promotes new scripts to production, while the software engineer develops and tests new features and functionalities of the core product, and the script developer uses the environment to develop and test the scripts. This process intensifies the collaboration among all the disciplines and promotes an agile way of working, making sure that RoboPlatform improves continuously.

Adding a new feature to RoboPlatform can take a few hours up to few months. To speed the script development process up, we have added the ability to create reusable components (called *metabots*) that the software developers can develop and that the script developers can use easily and quickly.

Recent Developments

The current generation of robots still are, for the most part, unintelligent virtual employees that dutifully follow rule-based scripts. RPA robots only work with bite-sized chunks of structured digital input. They cannot do much with a piece of paper or a pdf. We currently are making RoboPlatform smarter by letting it understand semistructured data, so the product can extract relevant data points for further processing. Self-learning classification algorithms are used; for example,



FIGURE 1. The robots within context.

a machine-learning model that simplifies plain unstructured text can be easily transformed into vectors and then classified by statistical methods, such as term frequency-inversed document frequency. In this way, we can dissect the required data points of items such as annual reports, salary slips, invoices, and income statements and then feed them to the robot in a structured file for further processing. We are enabling the use of the Python programming language to RoboPlatform, allowing developers to add data analytics to the scripts. We recently developed the complaints analyzer, a machine-learning module that supports customer advisors by prompting answers from past conversations. We are experimenting with natural language processing, to be able to interpret unstructured input such as legal contracts.

Talking about jobs in danger, we introduced the so-called blocks functionality in the newest Robo-Platform version: by simply dragging and dropping prebuilt metabots, users can build robot scripts themselves. For simple scripts, there is no need to ask a software developer to do the work anymore.

Taking Robots Global

ING decided to form a global RPA community two years ago. In this way, the bank can build toward global standards and global solutions for RPA. ING Slaski created the Robotics as a Service platform and the Robot Resource System in the bank's private cloud. The IT custodian and scripting teams can now provide RPA services globally. A good example of best of breed is integrating the RoboPlatform and a workload manager that ING Netherlands developed to manage the 3 million annual customer requests that 100 unattended robots process. This so-called Spider manages the workload based on service-level agreements, ensures secure storage of audit trails and log files, and enables robots and humans to interact. Whenever a robot is not able to finish a task that Spider assigned, the robot will pass the unfinished requests back to Spider, who then assigns the task to a human employee.

Beyond Banking

To remain competitive in software, volume is the key.⁵ Many businesses and applications can apply RPA.

ABOUT THE AUTHORS



PRZEMYSŁAW LEWICKI is a director of the Operational Digital Transformation Center and head of Robotics at ING Bank Śląski. Contact him at przemyslaw.lewicki@ingbank.pl.



JACEK TOCHOWICZ is a manager of the Robotic Process Automation Team at ING Bank Śląski. Contact him at jacek.tochowicz@ ingbank.pl.



JEROEN VAN GENUCHTEN is the global robotics product owner at ING Bank, The Netherlands. Contact him at jeroen.van .genuchten@ing.com.

As a bank, we can help customers with our knowledge of and experience with RPA and the product RoboPlatform that comes with it. We are now implementing a proof of concept with a number of corporate customers by allowing RoboPlatform to improve their own administrative processes. Applications are plentiful. We imagine expanding our robots' roles to be effective employees in other financial industries, medical applications, and public systems.

R ogue robots and robots building robots have inspired many Hollywood movies. Their impact on the economy and employment is a topic of attention in science and policy making.^{4,6} Here we have discussed one example of what is happening today and what we expect in the near future. The final question to answer is whether robots will take over our jobs. Robots will replace some jobs, and some new tasks will be assigned to robots from the start. At the same time, humans will be needed to build and control the robots. We believe that the combination of human beings and robots is the strongest model, in which employees can focus on customer-oriented activities while the robot does the mundane work. With the rise of smarter artificial intelligence algorithms, the portion of work that robots can do will increase, enabling humans to focus increasingly on important things. We hope employees can be trained to become business translators, supervising the algorithms and telling them *This article originally appeared in* IEEE Software, *vol. 36, no. 3, 2019.*

what data to use and how to interpret it. This robot model also applies to the software engineers and software: the end user will be more accustomed to building its own software using prebuilt blocks. However, software engineers will create the building blocks in years to come and, for the complicated scripts, we will need human intelligence in the foreseeable future. \mathbf{P}

References

- Gartner. Accessed on: Oct. 3, 2018. [Online]. Available: https://www .gartner.com/en/newsroom/press-relea ses/2018-10-03-gartner-says-roboticsto-become-mainstream-in-financedepartments-by-2020
- Everest Group, "Enterprise adoption of RPA exceeds 100% growth in 2017, buoyed by new buyers of all sizes, industries," Accessed on: Apr. 2018. [Online]. Available: https://markets .businessinsider.com/news/stocks /enterprise-adoption-of-rpa-exceeds-100-growth-in-2017-buoyed-bynew-buyers-of-all-sizes-industrieseverest-group-1027389360
- 3. C. L. Clair, A. Cullen, and M. King, "The RPA market will reach \$2.9 billion by 2021," Cambridge, MA: Forrester Research, Inc., Feb. 13, 2017. [Online]. Available: https:// www.forrester.com/report/The+RP A+Market+Will+Reach+29+Billion +By+2021/-/E-RES137229#
- 4. E. Brynjolfsson and A. McAfee, The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. New York: W.W. Norton, 2016.
- M. Genuchten and L. Hatton, "Software mileage," *IEEE Softw.*, vol. 28, no. 5, pp. 24–26, 2011.
- A.-F. Rutkowski, "Work substitution: A neo-Luddite look at software growth," *IEEE Softw.*, vol. 33, no. 3, pp. 101–104, 2016.

A Budding Romance: Finance and AI

Xiao-Ping (Steven) Zhang David Kedmey EidoSearch Enthusiasm for artificial intelligence and multimedia information in the financial industry is at an all time high. Every leader in finance now feels the pressure to answer the question, *"What is your AI strategy?"* Start-ups are playing a key role in helping the financial sector determine

what AI can do and how humans and machines can work together. In this essay, we describe emerging trends and attempts by FinTech start-ups to apply AI and multimedia information processing techniques across a wide range of business needs.

Enthusiasm for artificial intelligence in the financial industry is at an all-time high.¹ The trend began approximately ten years ago when a small cohort of startups sensed an opportunity to apply machine learning and multimedia processing to finance. It began with an explosion of digitized multimedia data and cheap computing power—driving forces that are still underway to this day.

These conditions provided fertile ground for entrepreneurs who imagined entirely new and automated workflows. They envisioned products that combined domain expertise in finance with knowledge of machine learning.

Awareness of this trend spread slowly, lurking under the radar, until grassroots activity crossed a threshold with the spontaneous arrival of its own name: "FinTech" (financial technology) entered our lexicon. The inaugural class of the FinTech Innovation Lab² in New York City in 2011 was an auspicious introduction. According to CBInsights, global FinTech deals in 2018 are on pace for a record year, reaching \$20.3 billion in VC-backed equity funding in the second quarter alone.³

In the last several years, a surge of interest in the promise of AI joined up with FinTech, kicking off a new phase of exponential growth. From a technologist's perspective, long-standing machine learning techniques were simply folded into the catchall phrase of AI. But then a major cultural shift occurred. The breakthrough successes of deep learning in multimedia, i.e., speech and image recognition, and AlphaGo, captured the imagination of people in finance and the general public.

Every leader in finance now feels the pressure to answer the question, "*What is your AI strategy*?" To provide answers, major resources are flowing to data science and AI groups within institutions and to the FinTech startups looking to serve them. Strategy depends on what aspect of the business you are addressing and the specific problem you are trying to solve. Financial institutions, like institutions in

any other industry, seek efficiencies in their operations: communications, customer service, human resource management, regulatory compliance, and fraud detection.

In this essay, we describe attempts by FinTech start-ups to automate aspects of business operations. We also address an intriguing problem that is in some ways unique to finance: the challenge of prediction in financial markets. The central challenge here—in a highly noisy and nonstationary system—is finding regularities in data that emerge from millions of traders and investors reacting to one another's decisions.

UNPACKING THE FINANCIAL INDUSTRY

Modern economies depend on a thriving financial sector, which in the U.S. accounts for 20 percent of GDP. Tens of thousands of banks, mutual funds, and hedge funds employ a vast army of financial professionals and technologists. This complex ecosystem serves one ultimate purpose that is to facilitate the flow of resources throughout the entire economy.

To understand the financial ecosystem, it helps to divide activities into two main functions. The brain of the ecosystem is an information-processing engine that detects where resources in the economy reside and where they should go. The body of the ecosystem is the infrastructure that supports this transfer. Below we organize the discussion by specific problems in operations (the body of the industry) and in making market predictions (the distributed brain) and the FinTech opportunities to address these challenges using AI.

A WORLD OF MULTIMEDIA SIGNALS

The financial industry continuously appraises signals in multimedia data: words spoken in audio and text recordings, consumer activity in the form of digital footprints, and economic dynamics from sensors and satellites that capture movement of goods and people. Processing this datum from raw form to actionable insight engages all of our multimedia techniques to help financial firms find relevant signals hiding in our midst. The proliferation of affordable drones and satellites is generating geospatial imagery that the company **Orbital Insight** processes to monitor and track global oil inventories—a key determinant of oil prices. Foursquare entered the arena of the Wall Street analysts several years ago, by translating its foot traffic data to accurately predict a new threshold of iPhone sales volumes. Foursquare was also able to anticipate plummeting sales at Chipotle.⁴ With daily checkins at 8 million locations—a total of 12 billion over the past nine years—traditional prognosticators must adopt new forms of data cleaning, automation, and analysis to keep up.⁵

Entirely new ecosystems of data vendors and data analyzers are emerging. **ExtractAlpha**, partnering with alternative data vendors such as **alpha-DNA**, processes data for use in systematic money management. Online consumer behavior information across multiple websites and search and social media platforms is compiled and organized by alpha-DNA in near real-time, and transformed by ExtractAlpha into The Digital Revenue Signal, a stock selection score designed to forecast revenue surprises based on changes in consumer demand. And, lest you think politics has avoided algorithmic scrutiny, new measures of company behavior in the environmental, social, and governance domain are being devised and modeled.

The hunt is on to find out what information these datasets hold to improve decision-making and which new sources of data can be brought into the fold. AI experts, as well as the multimedia techniques we have applied in so many other domains, must play a central role in finance.

MERGING HUMANS AND BOTS

AI, machine learning, and natural language processing cannot replace people—at least not yet. But humans and bots can work together in harmony to automate tedious tasks and enhance the human connection between business professionals and their customers. Let us consider three examples in the financial industry: wealth management, banking, and research.

Wealth management. The impact of AI in wealth management will not be wealth creation for the typical investor. Sorry to disappoint. In the article *To the Victor Go the Spoils: AI in Financial Markets*, it is argued from first principles that the "vast majority of people—no matter their level of AI expertise—will not achieve large excess returns."⁶ However, AI does bring efficiencies to operations. Wealth managers can serve more clients in less time and, in an increasingly digitized world, still deliver a personalized experience. To answer clients' questions and make informed recommendations, wealth managers must consider hundreds of data points: a client's investment portfolio holdings and transactions, trending market events, unexpected personal events, and evolving client needs. Start-ups such as **Forwardlane** address this need via AI, which prioritizes client Q&A from a natural language conversation interface, which helps wealth managers find precise answers to client questions.

Banking. Similar to wealth management, customers demand convenience and speed for banking services and loathe time spent searching for answers. How can banks offer these services at scale? A start-up at the forefront in this domain is **Kasisto**, which handles frequent banking tasks and, in a CoBot-like approach, knows when to hand-off to a live agent for services that need a human touch. The goal is to power "human-like conversations" through a conversational AI platform fluent in finance. Banks are eager to deploy these types of intelligent virtual assistants that can care for customers at a fraction of the cost and reduce call center volume.

Research. In order to assess the value of stocks, analysts scour reams of unstructured news and structured data that streams in from the web and accumulates within an organization. This human activity of processing data itself generates information, specifically, trails of search activity, discussions over email, and written reports. Machine intelligence algorithms that watch what a research analyst is reading and writing can track these unfolding events, offering fertile ground for Bots to learn and make recommendations to analysts of what to research. For example, the start-up **Diffeo** has collaborative agents that on their own might not know where to look, but by watching research teams do their work can help direct research efforts. One application is to discover entity connections between companies and people that emerge from such activity. Backed up by contextual evidence, these connections could be critical in assessing the likelihood of important events in the life of a company, for example, an impending merger or acquisition.

DETECTING BAD BEHAVIOR

Financial institutions are highly regulated and are expected to scrutinize the knowledge and intentions of their traders. The survival and success of a financial institution, regardless of size, comes down in large part to effective risk mitigation. This requires meaningful alerts and noise reduction for false alarms through automatic analysis of communications data. These efforts are costly and require technology to operate at scale. McKinsky & Company estimates that CEO fraud, where imposters gain access to business email accounts and fool unsuspecting employees to send funds to bogus accounts has led to losses of more than \$2.3 billion over the past few years.⁷

Business leaders are held responsible for the actions of their employees' actions. Take the following example given by the FinTech start-up, **Digital Reasoning**. When one trader messages "*Let's take this offline*" to another, is this an innocuous request to advance a meeting agenda, or an attempt to collude in secrecy? It depends on semantics and context. Using a private knowledge graph and natural language understanding technology, digital reasoning helps financial firms combine analysis of behaviors, intentions, and emotions. By transforming multimedia communications data, including text and audio, firms attempt to identify threats and mitigate reputational risk.

INSURANCE AND AI

Insurance companies have been in the business of evaluating risk for centuries. Yet many businesses in need of insurance find the process of obtaining quotes painfully inefficient. McKinsey & Company forecasts that "in 2030, manual underwriting ceases to exist for most personal and small-business products across life and property and casualty insurance."⁸ We are not there yet. Each year insurance underwriters receive applications from over 7.5 million small-medium sized businesses that fail to get an automatic quote. To solve this bottleneck, start-ups are ingesting data sources of a more granular

nature and applying machine learning techniques to permit analysis à la carte by geography and business classes. For example, **Open Data Nation** (ODN) aggregates billions of records published by city governments about commercial businesses and individual behaviors and builds machine learning models to anticipate issues. Insurance underwriters can then query ODN for an on-demand risk score better tailored to the unique risks of each business applicant.

USE MACHINES TO HIRE HUMANS

The most valuable asset of a financial firm is its workforce, so there is no more important process than the way a company attracts, selects, and retains talent. **Pymetrics** is a start-up that addresses this problem through the application of neuroscience games, which gather more insight about candidates and feeds this information through machine learning algorithms to increase the efficiency of hiring. More than ever before, financial institutions are in competition for tech talent. The firm **Untapt** is supplying AI algorithms to supplement human recruiters that review tech resumes. Using natural language processing, words on a resume are mapped relative to each other, each resume is sorted in a space with other resumes and their engine uses feedback to distinguish matches and nonmatches. By process millions of job scenarios, these algorithms learn to identify what qualities make for a successful candidate.

MARKET PREDICTION

The lifeblood of finance is information. Knowledge that provides more accurate and more *informative* predictions in financial markets translates directly into profit. Markets, however, are inherently noisy—much more so than traditional domains for AI. Changing relationships in markets pose unique challenges for AI researchers and, therefore, require unique solutions.

In addition to the inherent challenge of market prediction, data analysis needs have grown exponentially with the rise of alternative data. Investment banks and hedge funds are building new teams of data scientists to clean, structure, and analyze the fire hose of data streaming from every corner of the economy. Hundreds of alternative data start-ups are offering this new content. The list is long. It includes sentiment from social media, mobile data content, online reviews, and web searches; transaction data from e-commerce and credit cards; and new data from sensors such as satellite and geolocation data.

The basic question money managers are trying to answer is "what and how much can traditional and alternative data tell us about the future?" Also, an equally important question required to generate profit: "Is it unique? Does the information tell me something about the markets that others do not know?"

Helping to answer these questions, the start-up **EidoSearch** is a probability intelligence company that created a new type of AI to systematize the investment process and quantify prediction uncertainty.¹ Its numeric search engine finds conditions in data through a technique called data-incident based modeling, which takes advantage of multimedia signal processing and content-based retrieval technologies and is uniquely suited for nonstationary systems such as financial markets. Where deep learning and other forms of machine learning have fallen short, the EidoSearch method jettisons the need for a model with functional form. Instead, current events are automatically matched to similar data incidents, and their associated outcomes are used to generate a dynamic, model-free distribution forecast. This method, subjected to historical testing, has enabled hedge funds to detect new sources of profit. Each scenario tested is a "ProBot," which is a probability forecasting robot. Also, new evaluation measures for accuracy and informational uniqueness have been developed to select the most skilled (and profitable) forecasters among the ProBots.

PLANNING FOR THE UNFORESEEN

The growth of knowledge is inherently unpredictable, and awareness of what is unforeseen needs to be considered. The financial industry has great enthusiasm for AI and is wisely investing for the long term by bringing in experts from academia and the tech industry to help lead the way. These new hires are

critical points of contact for start-ups to market their AI services and to work collaboratively with in-house teams to both define and solve the pressing problems named in this essay. Using state-of-the art techniques is imperative. Two recent examples: Dr. Li Deng, a former chief scientist of AI at Microsoft, was recently hired as a Chief AI Officer at Citadel, one of the largest hedge funds in the world; and Dr. Manuela Veloso, on leave from Carnegie Mellon University, where she was the Head of the Machine Learning Department, recently joined the world's largest financial institution, JP Morgan Chase, to create and head an AI Research Center.

Dr. Deng is a leader in the speech recognition industry using large-scale deep learning—the successes of which served as a major impetus for the massive wave of interest in AI. Dr. Veloso, coming from the world of autonomous robots, is a particularly interesting and revealing choice. A CoBot, as her group defines it, is a robot that "follows a novel symbiotic autonomy, in which the robots are aware of their perceptual, physical, and reasoning limitations and proactively ask for help from humans."⁶

How is this relevant to finance? Replace the concept of navigating the physical world with the demands of navigating information flows within finance, i.e., a "virtual" overlay of our physical economy. New paths ahead in the world of finance are being forged—in concurrence with creative start-ups—as humans and machines learn to work symbiotically.

REFERENCES

- M. Kolanovic and R. T. Krishnamachari, "Big data and AI strategies: Machine learning and alternative data approach to investing," JP Morgan Chase & Co., New York, NY, USA, Tech. Rep., May 2017.
- 2. FinTech Innovation Lab. [Online]. Available: http://www.fintechinnovationlab.com/
- 3. CBInsights, "Global Fintech report Q2 2018," New York, NY, USA, 2018.
- 4. J. Wieczner, "Foursquare just predicted Chipotle's sales Will plummet 30%," Fortune, Apr. 15, 2016.
- 5. E. Stinson, "Foursquare may have grown up, but the check-in still matters," Wired, Aug. 9, 2017.
- 6. X.-P. Zhang, "To the victor go the spoils: AI in financial markets," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 171–176, Nov. 2017.
- J. Corbo, C. Giovine, and C. Wigley, "Applying analytics in financial institutions' fight against fraud," McKinsey & Company, New York, NY, USA, Apr. 2017. [Online]. Available: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/applyinganalytics-in-financial-institutions-fight-against-fraud.
- R. Balasubramanian, A. Libarikian, and D. McElhaney, "Insurance 2030—The impact of AI on the future of insurance," McKinsey & Company, New York, NY, USA, Apr. 2018. [Online]. Available: https://www.mckinsey.com/industries/financial-services/our-insights/ insurance-2030-the-impact-of-ai-on-the-future-of-insurance.
- 9. CoBot Robots. [Online]. Available: http://www.cs.cmu.edu/~coral/projects/cobot/

ABOUT THE AUTHORS

Xiao-Ping (Steven) Zhang is a Co-Founder of EidoSearch. Contact him at xzhang@eidosearch.com.

David Kedmey is a Co-Founder of EidoSearch. Contact him at dkedmey@eidosearch. com.

This article originally appeared in IEEE MultiMedia, *vol. 25, no. 4, 2018.*

Project-Based Learning Inspires Cybersecurity Students: A Scholarship-for-Service Research Study

Alan T. Sherman | University of Maryland, Baltimore County Peter A.H. Peterson | University of Minnesota Duluth Enis Golaszewski, Edward LaFemina, and Ethan Goldschen | University of Maryland, Baltimore County Mohammed Khan | Prince George's Community College Lauren Mundy | Montgomery College Mykah Rather | Prince George's Community College Bryan Solis | Montgomery College Wubnyonga Tete | Prince George's Community College Edwin Valdez | Montgomery College Brian Weber and Damian Doyle | University of Maryland, Baltimore County Casey O'Brien | Prince George's Community College Linda Oliva | University of Maryland, Baltimore County Joseph Roundy | Montgomery College Jack Suess | University of Maryland, Baltimore County

> ver four summer days in 2017, cybersecurity students at the University of Maryland, Baltimore County (UMBC) analyzed the security of a targeted portion of the UMBC campus network, discovering numerous flaws, creating proof-of-concept exploits, and providing practical recommendations for mitigation. We report on this novel summer research study; its technical findings; and takeaways for students, educators, and Information Technology Departments.

> UMBC, a National Center of Academic Excellence in Cyberdefense Education and Research, is a midsize public university offering undergraduate and graduate tracks in cybersecurity leading to B.S., M.S., and Ph.D. degrees in computer science, computer engineering,

Digital Object Identifier 10.1109/MSEC.2019.2900595 Date of publication: 14 May 2019 and information systems and the master of professional studies degree in cybersecurity. UMBC is also a Cybercorps: Scholarship for Service (SFS) school, where students are supported for up to three years on the condition that, after graduation, they will work for federal, state, local, or tribal governments one year for each year of support.

In the fall of 2016, with support from the National Science Foundation, UMBC was one of 10 schools that pioneered a new strategy for recruiting talented cybersecurity professionals for government service: the university extended SFS scholarships to nearby partnering community colleges (CCs). To integrate the new CC students into the existing SFS cohort through a collaborative activity, Alan T. Sherman, UMBC professor and director of UMBC's Center for Information Security and Assurance (and one of the authors of this article), organized a four-day SFS summer research study at UMBC in the summer of 2017. Prof. Sherman also invited professors, researchers, UMBC graduate students, and National Security Agency (NSA) personnel to interact with the students as technical experts.

Everyone worked as a team on the same challenge: to analyze the network administration system's (NetAdmin's) web front end enabling modifications to the UMBC campus firewall. In support of the project, UMBC's Division of IT (DoIT) provided participants with all relevant source code and a functional copy of the environment for testing. At the end of each day, DoIT staff, including the primary NetAdmin author, met with the students. At the conclusion of the project, the student team identified several critical

26

vulnerabilities, devised exploits, and presented their findings and recommendations to DoIT.

This type of activity should be beneficial for any group of students. Our hope is that educators, IT Departments, and students at any institution may learn from our shared experiences in collaborative and real-world

project-based learning (PBL) (see "Project-Based Learning"). Partnering with a real IT Department has many benefits: the study inspired students and enhanced students' skills, students and educators appreciated the authentic case study, DoIT received free security consulting, and the UMBC community gained improved security.⁵

The SFS Summer Study at UMBC

A hands-on study was appealing because it enabled collaboration,

problem solving, and independent thinking in addressing an important, practical, rich, and challenging

Our task was to analyze the security of NetAdmin and the network architecture and to make recommendations to DoIT.

problem. We sought a problem that was complex but tractable. We also sought a project that, if successful, would benefit the UMBC community. Focusing on UMBC's home-grown NetAdmin had many attractive properties: NetAdmin's source code was available; DoIT could answer questions and provide information; and, since NetAdmin had never undergone a security evaluation, it seemed likely to have vulnerabilities.

The in-person participants comprised six CC transfer students, three UMBC undergraduates, and one Ph.D. student. All students had at least a basic grounding in cybersecurity. Some students had much

> more expertise. Each participant signed a nondisclosure agreement (NDA) with DoIT.

> The study took place from 9 a.m. to 5 p.m., Tuesday through Friday, in a large room with tables, a whiteboard, and

a projector. Using a PBL approach,² we presented the challenge and challenge-related goals to the students and instructed them to formulate a strategy that would achieve the project's goals, while supporting sustained inquiry and reflection. Students organized themselves into teams, with each team exploring some aspect of the problem. For example, teams explored the network topology, the software environment, architectural issues, source code, and known software vulnerabilities. More experienced students emerged as leaders.

Project-Based Learning

P roject-based learning (PBL) is an instructional approach in which small groups of students engage in authentic tasks and learn as they attempt to solve relevant problems. Students ask and revise questions, debate ideas, generate predictions, experiment, collect data, draw conclusions, communicate ideas and findings, refine approaches, and create products.²

PBL holds great promise in cybersecurity because there is a proliferation of complex challenges to engage students, sustain their interest, and direct their learning as they develop diverse approaches to solving real-world problems. In PBL, students are focused on tasks; they can try out a variety of solutions and receive timely feed-back on their approaches. They engage in collaboration and reflection that deepens their learning and enhances the transferability of skills.

There are many examples of PBL in cybersecurity (e.g., the New Jersey Institute of Technology's Cyber-Real World Connections Summer Camp^{S1} and Conklin and White's graduate course, ^{S2} which includes some elements similar to our study). We encourage the creation of more scholarly articles on this subject. We are strong believers in the value of PBL, as evidenced by our participation in the INSuRE Project. ^{S3}

References

- S1. New Jersey Institute of Technology, "Cybersecurity Real World Connections Summer Boot Camp at NJIT," 2016.
 [Online]. Available: https://sci.njit.edu/gencyber/RWCCybersecurityCamp Brochure-Summer2016.pdf
- S2. A. Conklin and G. White. "A graduate level assessment course: A model for safe vulnerability assessment," in *Proc. 9th Colloquium for Information Systems Security Education (CISSE)*, June 2005, pp. 109–114.
- S3. A. T. Sherman et al., "The INSuRE Project: CAE-Rs collaborate to engage students in cybersecurity research," *IEEE Security Privacy*, vol. 15, no. 4, pp. 72–78, July–Aug. 2017.



Figure 1. An illustration showing the architecture of the UMBC network, including the NetAdmin tool, which is accessible to machines on the research subnet.

Two UMBC professors and two NSA experts visited each day to answer technical questions. Late each afternoon, representatives from DoIT, including the primary NetAdmin script author, joined the group for a discussion. Students unable to attend in person joined a student-led one-hour evening chat session via Google Hangouts.

The Problem

The UMBC network has 10,000 users; more than 15,000 devices connect to the network daily. That

makes defending the UMBC network a daunting challenge. One part of the defense is a firewall between the Internet and the UMBC network. All campus traffic must pass through this firewall. One of UMBC's inter-

nal subnets is for computers used in research projects. Users on these computers often need to connect to and from the Internet on various ports. This requires permission to enable data to pass through the firewall. DoIT originally processed firewall exceptions manually, which was time-consuming and error prone. NetAdmin, launched in 2006, facilitates exceptions to UMBC's default-deny firewall policy. Our task was to analyze the security of NetAdmin and the network architecture and to make recommendations to DoIT.

NetAdmin allows faculty and staff who are authenticated through the myUMBC single sign-on (SSO) system to create firewall exceptions for their machines on the research subnet. As shown in Figure 1, NetAdmin sits behind the UMBC firewall, so it can be accessed only from the campus network or by virtual private network (VPN) users.

The adversary's main goal was to make unauthorized changes to the UMBC firewall without detection.

User groups, including faculty, staff, and superusers, are defined in a file in NetAdmin's application directory. Superusers may view, modify, or create any rule for any Internet Protocol address on the UMBC network (not only on the research subnet). Faculty and staff may create, modify, or delete rules for certain common ports [e.g., Secure Shell (22), HTTP (80)] associated with research subnet addresses they "own." Rules violating these restrictions must be submitted out of band to DoIT for special consideration. Since machine owners could modify only rules affecting their own machines, DoIT reasoned that Net-Admin introduced little risk.

Written in PHP 5.1.6 and residing on a dedicated Linux server running Apache 2.2.3, NetAdmin receives firewall rules from client browsers and applies those rules to UMBC's firewall through application programming interface (API) calls. To authenticate the rules to

> the firewall, NetAdmin includes a 360-bit symmetric API key file stored in the application directory of the NetAdmin server. This file is neither digitally signed nor integrity protected. In case of failures and

restarts, NetAdmin stores rules and logs in local unstructured files. Each rule is described by one record, which is delimited by a newline. Pipe characters delimit fields.

For more than a decade, NetAdmin ran untouched and worked well, with no detected compromises. No one, however, had ever subjected NetAdmin to a thorough security evaluation. In planning discussions, DoIT suggested analyzing NetAdmin in the same way that a penetration testing team might. Students were encouraged to follow whatever approach they thought best and were given access to DoIT staff, who provided appropriate information as requested.

Our adversarial model was an outsider with compromised faculty or staff credentials or a malicious faculty or staff insider on the research subnet with the knowledge, skills, and resources of an excellent computer science graduate student. The adversary's main goal was to make unauthorized changes to the UMBC firewall without detection. The group analyzed NetAdmin in its operational context, including whether cryptography was being properly used, but did not consider attacks on the cryptography itself, the servers' physical security, social engineering of DoIT staff, or recovery after disaster or compromise.

Vulnerabilities, Attacks, and Risks

At the start of our four-day study, the student-led team of 10 individuals focused on identifying risks, potential vulnerabilities, and related attacks,

many of which were extremely serious. NetAdmin ran on an unpatched, out-of-date, and unsupported operating system (OS), Linux 2.6.18, which has at least 463 vulnerabilities (https://www .cvedetails.com). Violating the principle of least privi-

lege,¹ the firewall API key used by Net-Admin permitted arbitrary changes to the campus firewall (not just to the research subnet). Compromise of the NetAdmin server would therefore be very severe. An attacker could issue arbitrary firewall rules affecting the entire campus; modify log files, rules, and user groups; and exfiltrate the firewall API key, all of which are stored as unencrypted text without integrity protection.





Students found some of the most common software security errors.³ NetAdmin did not adequately validate or sanitize inputs. For example, NetAdmin permitted firewall rules to include text descriptions but did not strip HTML or JavaScript. This made it possible for someone to conduct code injection attacks,⁴ which could victimize users

While DoIT was not aware of any attack involving NetAdmin, the potential attacks listed were feasible and could be executed by skilled students.

and administrators through their browsers. JavaScript payloads could submit rules to NetAdmin in the background. The malicious code could execute arbitrary commands on the NetAdmin server. The malicious code could, for example, initiate commands to exfiltrate the firewall API key.

Similarly, NetAdmin did not validate the length of rule descriptions, which meant that the system could be vulnerable to possible record-overflow attacks and/ or denial-of-service attacks. In particular, NetAdmin's use of the PHP command **fgetcsv()** assumed (without verifying) that each record was at most 999 bytes. As shown in Figure 2, if a user (or adversary) entered a rule longer than 999 bytes, the additional bytes would be accepted as a

new and valid record.

Communication between users and NetAdmin was unencrypted HTTP without integrity protection, allowing an adversary to read and modify all traffic. By modifying data

sent to NetAdmin, an adversary could set firewall rules enabling unauthorized access to the user's machines or launch an injection or record-overflow attack. Also, while NetAdmin authenticated the firewall using a self-signed certificate, the firewall did not authenticate NetAdmin; it required only that requests contain the API key. Additionally, since the firewall's key was self-signed, compromise of



Figure 3. An illustration showing the recommended architecture to provide compartmentalized defense. This design restricts failure of the research subnet firewall to the research subnet.

UMBC's signing key could enable an adversary to forge certificates and impersonate the firewall.

Other risks were exposed. For example, UMBC's one-firewall design provided no architectural protection. NetAdmin was accessible via the campus VPN, facilitating remote attacks. If an adversary could hijack a user's SSO session, that adversary could masquerade as that user to NetAdmin.

While DoIT was not aware of any attack involving NetAdmin, the potential attacks listed were feasible and could be executed by skilled students. As proof of concept, students implemented record-overflow and injection attacks.⁵

Recommendations

After identifying attacks, the students recommended a number of mitigations: the NetAdmin software, including the OS and all supporting software, should be kept current with security patches to mitigate off-the-shelf exploits; all input should be sanitized and validated on the server side; HTML, Javascript, and special characters (e.g., pipe) should be prohibited in rules; and size limits should be enforced to stop overflow attacks.

Also, NetAdmin should use different API keys for superusers and

faculty, with the latter affecting only the research subnet. API key establishment and storage might be improved by encrypting the API keys and keeping digests for integrity checking. The digests could be kept offline for periodic manual integrity checks, but the plaintext API keys are actively needed by the server during operation; keeping the encrypted API keys and digests locally would have limited value given that there is no secure place on the NetAdmin server to store them. As mentioned, compromise of the NetAdmin server would be catastrophic; in this case, the keys would be revealed. There is no perfect solution for the key-storage issue.

Figure 3 shows a two-firewall approach with better segmentation, where the research subnet firewall and the main campus firewall use separate keys. Regardless, communications between the NetAdmin server and users should use end-to-end encryption with authentication and integrity protection, and the firewall and NetAdmin should authenticate each other using certificates signed by a certificate authority.

Using a direct, physical connection between NetAdmin and the proposed research subnet firewall would improve physical

security. Segmenting NetAdmin into a web front end, for validating and sanitizing input, and a back end, for performing additional validation and for communicating with the firewall, would add defense in depth. These services should run under separate accounts and be restricted in other ways (e.g., no unnecessary software or communication with unnecessary hosts). Disallowing connections from the campus VPN would reduce the potential for remote attacks, though it would be difficult to prevent an adversary from logging into NetAdmin after establishing a VPN connection to another campus machine. Performing periodic internal and external audits of NetAdmin's software and firewall rules would help sustain security.

Takeaways

We hope that educators, IT departments, and cybersecurity program managers can benefit from our experience.

Educators and Study Organizers

Overall, the study went very smoothly, and PBL sustained inquiry and critical thinking. Most students quickly became absorbed in the project and were productive, although some students could have benefited from some prior preparation. Engagement level varied, but everyone made contributions. A few students were somewhat uncomfortable with the undirected and open-ended model. However, in a follow-up survey, 100% of participants reported that the project increased their cybersecurity knowledge and skills (86% strongly agreed and 14% agreed). Participants identified the following elements as valuable: teamwork, hands-on nature of the task, real-world challenge, critical thinking, and problem solving. All participants reported that they would recommend the summer study project to other cybersecurity students.

Having a virtual copy of the production system for experimentation was extremely valuable as was having access to the original developer. Posing questions to DoIT in a Google Doc and receiving answers throughout the day was effective and helpful as was having local security experts available for consultation. In-person discussions were facilitated by a video projector, whiteboard, and students' personal devices. We recommend having numerous power strips available. Evening chat sessions allowed remote students to participate. Chat worked better than video because it provided a written record and facilitated asynchronous use. Summer internships can create scheduling conflicts; we now hold the study during the January intersession.

IT Departments

IT departments often run obsolete and unpatched systems because they know that updates will take valuable staff time and might break the system, requiring even more staff time to fix. Our study, however, demonstrates that keeping software systems up to date is not optional. We also exposed and exploited numerous common vulnerabilities and suggested improvements. IT departments elsewhere could benefit from similar analysis.

We were fortunate to enjoy remarkably strong support and cooperation from DoIT, and we commend members of the department for their constructive attitude. Teams at other schools, however, might face a defensive administration that fears embarrassment or is unwilling to trust students. We believe that careful selection of participants and the use of NDAs should reassure administrators that students in the project can be trusted. Our hope is that, by welcoming and encouraging analysis of their systems, other IT departments and student teams can learn while enhancing the security of their communities.

Cybersecurity Program Managers

Extending scholarships to CC students has thus far has worked well. In recruiting CC students for our SFS program, we focus primarily on those pursuing associate degrees because they are more prepared to transfer to four-year schools, even though some associate of applied science programs include more cybersecurity coursework. While there is an opportunity cost in that a scholarship awarded to a CC student is not awarded to a student at UMBC, we are attracting highly qualified CC students, and the scholarship is a life-changing opportunity for some students, especially those from modest backgrounds. Our current approach is to support two CC graduates per year.

O ur study engaged and motivated students, as evidenced by their findings and our survey results. We also demonstrated that there are highly capable students at CCs who can contribute to cybersecurity. While we integrated this study into the SFS program at UMBC, we feel this type of activity

could be integrated into nearly any kind of cybersecurity program. Partnering qualified students with IT Departments can reap benefits for everyone: students gain exciting, concrete, hands-on collaborative experiences; educators are given rich and realistic case studies supporting project-based learning; and IT Departments receive free cybersecurity consultations. DoIT hired several of the participants to join its security team. We look forward to conducting similar studies each year and hope that other schools can also benefit from similar collaborations.

References

- M. Bishop, Computer Security: Art and Science. Boston: Addison-Wesley, 2003.
- P. C. Blumenfeld, E. Soloway, R. W. Marx, J. S. Krajcik, M. Guzdial, and A. Palincsar, "Motivating project-based learning: Sustaining the doing, supporting the learning," *Educ. Psychol.*, vol. 26, pp. 369–398, 1991.
- S. Kaza, B. Taylor, and E. K. Hawthorne, "Introducing secure coding in CS0, CS1, and CS2: Conference workshop," J. Computing Sci. Colleges, vol. 3, pp. 11–12, June 2015.
- Open Web Application Security Project, "The OWASP Foundation: The free and open software security community." [Online]. Accessed on: Sept. 16, 2018. Available: https://www.owasp.org/index .php/Main_Page.
- A. T. Sherman et al., The SFS Summer Research Study at UMBC: Project-based learning inspires cybersecurity students, *Cryptologia*, to be published, Nov. 2018. [Online]. Available: arXiv:1811.04794.
- Alan T. Sherman is with the University of Maryland, Baltimore County. Contact him at sherman@ umbc.edu.
- **Peter A.H. Peterson** is with the University of Minnesota Duluth. Contact him at pahp@d.umn.edu.

- Enis Golaszewski is with the University of Maryland, Baltimore County. Contact him at golaszewski@ umbc.edu.
- Edward LaFemina is with the University of Maryland, Baltimore County. Contact him at edlafem1@umbc .edu.
- **Ethan Goldschen** is with the University of Maryland, Baltimore County. Contact him at egold2@ umbc.edu.
- Mohammed Khan is with Prince George's Community College. Contact him at khanmoh1@ umbc.edu.
- Lauren Mundy is with Montgomery College. Contact her at lmundy1@umbc.edu.

- Mykah Rather is with Prince George's Community College. Contact her at mrather1@umbc.edu.
- Bryan Solis is with Montgomery College. Contact him at bsolis1@ umbc.edu.
- Wubnyonga Tete is with Prince George's Community College. Contact her at wtete1@umbc.edu.
- Edwin Valdez is with Montgomery College. Contact him at evaldez2@umbc.edu.
- Brian Weber is with the University of Maryland, Baltimore County. Contact him at brianw5@umbc.edu.
- Damian Doyle is with the University of Maryland, Baltimore County. Contact him at damian@umbc.edu.

- **Casey O'Brien** is with Prince George's Community College. Contact him at cobrien@nationalcyberwatch .org.
- Linda Oliva is with the University of Maryland, Baltimore County. Contact her at oliva@umbc.edu.
- Joseph Roundy is with Montgomery College. Contact him at Joseph. Roundy@montgomerycollege.edu.
- Jack Suess is with the University of Maryland, Baltimore County. Contact him at jack@umbc.edu.

This article originally appeared in IEEE Security & Privacy, vol. 17, no. 3, 2019.



IEEE Security & Privacy magazine provides articles with both a practical and research bent by the top thinkers in the field.

- stay current on the latest security tools and theories and gain invaluable practical and research knowledge,
- learn more about the latest techniques and cutting-edge technology, and
- discover case studies, tutorials, columns, and in-depth interviews and podcasts for the information security industry.



computer.org/security

Digital Object Identifier 10.1109/MSEC.2019.2911826

Internet Computing

IEEE Internet Computing delivers novel content from academic and industry experts on the latest developments and key trends in Internet technologies and applications.

Written by and for both users and developers, the bimonthly magazine covers a wide range of topics, including:

- Applications
- Architectures
- Big data analytics
- Cloud and edge computing
- Information management
- Middleware
- Security and privacy
- Standards
- And much more

In addition to peer-reviewed articles, *IEEE Internet Computing* features industry reports, surveys, tutorials, columns, and news.

www.computer.org/internet



Join the IEEE Computer Society for subscription discounts today!

www.computer.org/product/magazines/internet-computing



Department

Learning to Network

Stephen D. Crocker Shinkuro, Inc.

Editor: David Walden [dave@walden-family.com]

THE ARPANET FOR me was an intriguing diversion from my graduate studies for the summer of 1968. During the previous eighteen months, I had studied artificial intelligence (AI) at Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. I decided to spend the summer at my undergraduate school, University of California, Los Angeles (UCLA), CA, USA, to work with both Professor Gerald (Gerry) Estrin and my friend Vint Cerf. I packed my bags, sublet my apartment in Cambridge, and headed west.

Shortly after arriving at UCLA, Vint described an unusual graduate meeting he would attend in late June. Rather than a conventional conference with papers and a common research theme, the sole purpose of the meeting was to introduce graduate students who were working on projects supported by the Information Processing Techniques Office (IPTO) in the Advanced Research Projects Agency (ARPA).

I was mildly jealous but I was not a UCLA graduate student and had no right to complain. Luckily, IPTO also funded research at MIT, including Marvin Minsky's AI Laboratory. Minsky selected Pat Winston and me to represent his lab, and his secretary tracked me down. The conference in Illinois was attended by approximately 30 guys, one or two from each IPTO funded project. A networking meeting in the social sense, it was

Digital Object Identifier 10.1109/MAHC.2019.2909848 Date of current version 29 May 2019. an opportunity for graduate students from the various projects to meet others and share experiences. The IPTO principal investigators met every year, and this was an experimental spin-off at the graduate student level.

Barry Wessler, the sole program manager in IPTO, organized and ran the meeting. Barry was warm, welcoming, and roughly the same age as we were. He structured the meeting for the group to become acquainted with each other, reserving only one block of time for a discussion of computer networking. For most of the three days, I listened to descriptions of current cutting-edge computer science research: advanced interactive systems; AI; graphics; multiprocessors; new programming languages; etc. One of the more engaging presentations covered the possibility of small portable computers. It was a fantastic idea for an age in which computers were large, bulky, and filled air-conditioned rooms. Alan Kay, then a graduate student at the University of Utah held forth: "And by 'portable'," he proclaimed, "I mean you can carry it and something else as well."

During the discussion of computer networking, Barry described the ideas that would eventually become the Arpanet. It did not capture the group's attention; we were each focused on our own research goals. I doubt many of us were aware that IPTO was fully committed to building the network which would directly affect each of us. The IPTO people, Larry Roberts, Bob Taylor, and Barry, had been working on the plan for the network, developing the technical strategy while also gaining the acceptance and cooperation of the principal investigators at each of the research sites.

The Arpanet plan was already well advanced. The initial four sites would be in the western part of the United States at UCLA, Stanford Research Institute (SRI), Menlo Park, CA, the University of California, Santa Barbara (UCSB), CA, and the University of Utah, Salt Lake City, UT, USA. Wes Clark's suggestion to place a separate computer at each site to perform the packet-switching instead of requiring the host computers to do this work was already part of the plan. These were dubbed "IMPs" for interface message processor. This was the birth of routers, the central element in all successive network architectures.

IPTO had SRI organize a meeting with the institutions to discuss the basic elements of the network. The meeting was held at UCSB in August; Vint and I attended on behalf of UCLA.

I had known Vint since high school where we pursued our common interest in math and began to get involved in computing. While still in high school, I was given access to some of the UCLA computers and I occasionally dragged Vint along on some visits. We attended different colleges, he at Stanford and I at UCLA. Vint then attended graduate school at UCLA while I went off to MIT. The summer was a fortunate chance to work together.

Elmer Shapiro from SRI chaired the meeting at UCSB. Shapiro developed a preliminary description of the new network. There were roughly a dozen of us in the room, two from each of the four sites and a few observers.

We represented the second level of the research personnel. We were graduate students or staff members. None of the principal investigators attended. Although we were not in charge of the research agendas in our home institutions, we had a thorough understanding of the computers and software, including operating systems, compilers, and the applications. This common base would be crucial in figuring out how to connect our four different computer systems to form the first nodes of the network.

The meeting was not highly structured. Shapiro explained the basic plan for the project and offered a bare outline of the Arpanet architecture. Our sites would be connected by long-distance 50 kb/s lines. The lines would be shared using a technology called packet switching. The lines would be connected to our computers via the IMPs. ARPA was in the process of soliciting bids to build the IMPs.

Shapiro had come prepared with possible experiments using dial-up lash-ups between our machines before the network was delivered. However, we found ourselves more interested in the challenges of connecting our computers with the new technology. The details of the hardware connections would depend on the design of the IMP and how the contractor built it. However, to make those connections work, we would have to make incisions into both the hardware and the operating system on our computers, and we would have to design the kinds of messages that our computers would say to each other. Our four laboratories had machines that were commonly used in computer science research but they were of four different designs. UCLA had a Scientific Data Systems (SDS) Sigma 7 computer, SRI had an SDS 940, Utah had a Digital Equipment Corporation PDP-10, and UCSB had an IBM 360/75. The machines had different operating systems and different character sets. All had the capacity for timesharing, but each behaved as if it were the center of its own world and had no easy way of connecting to another machine.

Our group did not make technical decisions at that meeting, but we realized we needed to continue talking. We did not explicitly think of ourselves as an organization, per se, but we implicitly gravitated toward forming our own human network. We needed to understand each other's computing environment and we agreed to hold meetings at each of our sites. In making this decision, we recognized the implicit irony. We would do a substantial amount of travel in order to build a network that would eventually permit collaboration without travel.

The group from the Santa Barbara meeting became the nucleus of a working group that cut across communities and bound the laboratories together. It grew from an informal group of fewer than a dozen to what is now the Internet Engineering Task Force (IETF). Fifty years later, this task force engages thousands of people across the globe who continue to develop hundreds of protocols.

My involvement in the Arpanet was supposed to last the summer. As time neared for me to return to Cambridge, I realized I was not anxious to leave. I shared my thoughts with my boss for the summer, Gerry Estrin. He immediately invited me to join the UCLA Ph.D. program. I debated the change since I had chosen to do my graduate research in AI, where MIT was a clear leader. After a few sleepless nights, I chose UCLA, returned to Cambridge to withdraw from MIT and close my apartment.

The fall of 1968 was a time of travel. Our informal group visited each of the four ARPA sites, talked about the local research agendas and brainstormed ideas on what our computers would say to each other. Our discussions remained general; we did not yet have the concrete specifications for the network. Rather than focusing on the details of message formats, exact sequences of messages, etc., we explored more general concepts. We knew we wanted to build a network that would handle the unique and creative elements of any machine or system without forcing those elements into a homogenized or restricted form, so our thinking focused on generality instead of minimal functionality

Generality is fine as a principle, but it does not provide much guidance. Meanwhile, although we did not want to close the door on potentially interesting uses of the net, we could see two fundamental services that would be immediately useful. First, users at each site would like to login to distant machines just as if they were connecting to that machine over conventional telephone lines or directly connected terminals. Second, users would want to be able to transfer files from one machine to another. At the same time, we did not want to focus only on these two services lest we compromise the larger possibilities.

Our thinking included research systems beyond those at the first four nodes. We were all aware of Multics at MIT and how it represented the next generation in time-sharing systems. We all knew that ILLIAC IV at the University of Illinois, IL, USA, was planned to be a big step forward in parallel computing that would be able to handle large numerical calculations. We were also interested in interactive natural language systems, chess programs, databases, graphics, and more. Beyond the existing systems, we wanted to create the environment for applications *that could use the facilities at two or more sites at the same time*. This meant we had to be careful not to build unnecessary assumptions or constraints into the basic infrastructure.

Two of the examples influencing our thinking about how the network should interact with existing systems were the Culler-Fried System at UCSB and Douglas Engelbart's On-Line System (NLS) as SRI. The Culler-Fried System was a clever tool for experimenting with signal processing. Like a desk calculator on steroids, single button pushes caused computations, but it operated on vectors of 128 numbers and reduced complicated operations, such as convolution, to a single keystroke. Douglas Engelbart's NLS was the precursor of the graphical user interface for text. It had a mouse and allowed users to interact with structured text and hyperlinks. To support these over the network would be challenging. At the very least, we knew it would be important to transmit single button pushes if the remote system required it, even though that might seem inefficient. In other cases, it would be important to move larger chunks of data as efficiently as possible.

During our discussions that fall, we identified two principles that would guide our work. First, our network services or protocols were to be constructed in layers and these layers would be as thin and as simple as possible. Our second principle was the entire network structure would be open. The different layers of protocols needed to be accessible to any user. All users should be able to modify the protocols, add new layers, and insert new protocols between layers.

In hindsight, our decisions felt natural to us because they reflected the cooperative research environment in which we worked. Hence, it made sense to have the internal structure open and available to the users. Had we been building a commercial product, it would have been natural to have a hard boundary between that internal structure and the services that were available to the users.

Although we did not spend time saying so explicitly, we were conscious that we lacked any formal authority for designing the structure of network software or for defining network standards. We were simply graduate students who might develop some of the network software. However, we had been visiting the initial sites and had built up an understanding of the requirements for such a system. None of us rushed to exert our authority over network software. In those early days, we met as equals and without hierarchy.

February In 1969, our environment expanded when our group began working with Bolt Beranek and Newman, Inc. (BBN), the contractor that had been selected to build the IMPs. The BBN development team was led by Frank Heart. Frank was a seasoned, no nonsense digital engineer. He started his career in the early 1950s working on the Whirlwind at MIT. He then moved to Lincoln Labs and worked on the communications circuits for the SAGE computers, the machines the Air Force used to monitor intrusions into the U.S. airspace. He would later claim that his group at BBN knew more about computer communications "than any other group in the country."¹

The IMPs were based on a commercial product, the Honeywell 516 computer, that Frank's group augmented for the Arpanet project. His engineers specified modifications for the Honeywell hardware. They also designed and wrote the software that performed the IMP functions.

Frank and his team were accustomed to working in an industrial environment, where they met deadlines and contract specifications. They were aiming to deliver the IMPs on a very tight schedule. The first IMP was scheduled for UCLA in September, and the others on a monthly schedule. Although they were building a research network, they expected the network would be reliable and always available.

On a snowy day that February, the two groups, West Coast graduate students and East Coast engineers, met at the BBN offices. The two groups approached the technical problems from different perspectives. Frank's team had thought about the details of the IMP and how it would operate. We had spent our time thinking about how our computers, which were called the hosts, would talk with the IMPs. As we considered potential solutions, we also recognized we needed to deal with issues that were organizational or bureaucratic as well as technical. At this initial meeting, we sized up each other and tested our future relationship. We had come prepared with some preliminary ideas. One of our members, Jeff Rulifson of SRI, suggested good practice would be to insert, at various layers, some lightweight checksums for data integrity. He had argued this approach provided substantial benefits in operating systems by catching software and configuration errors.

We shared our plan to use a simple checksum to catch major errors, including the possibility of incorrectly assembled messages in the IMPs. Frank Heart pushed back very forcefully, booming, "You'll make my network look slow."

I was uncertain how to respond. The lines of authority were far from clear. It really was not for him to tell us how to design our protocols but we had no formal authority. I did not want our relationship to start in a contentious manner. I focused on the planned bit-serial host-IMP interface and tried to make a point about potential errors. "How reliable is that interface?" I asked. "As reliable as your accumulator!" Heart boomed again. The accumulator was the key element of a mid-1960s computer. If it failed, your computer was broken. His vehemence convinced us to remove a checksum from our plans. This was a mistake. When Lincoln Laboratory added their TX-2 computer to the network in 1970, they had a lot of trouble debugging their software. It turned out their hardware interface to the IMP had some crosstalk with their drum storage unit. Rulifson's advice would have paid off.

Our meetings with Heart and his BBN staff were productive, and we left on good terms. Yet, we returned to our home institutions with the feeling that we were in an awkward position. Although we were making decisions about network design and implementation, we had no formal authority. We realized we needed to put our ideas on paper. Even though we were meeting every few weeks, we lacked enough time to discuss all our ideas. We also needed to include others in our thinking. When we met a month later at the University of Utah, we agreed to start documenting our ideas.

We each took a writing task. I took on the additional clerical assignment of organizing the notes, which seemed to be a minor chore. However, each time I started to jot down how we might organize these notes, I felt blocked. I feared the simple act of writing these notes might trigger a backlash from someone in charge—someone from the East, maybe Boston or maybe Washington—asking who we thought we were.

Our group was deeply involved in the design and implementation of the network. We had insights, both practical and theoretical, that would help all of us build this technology. If we neglected to capture our thoughts in writing, we would be retreating from our assignment to develop this network. I found myself struggling with this problem evening after evening. Finally, I realized one of the key lessons of networking, i.e., you have to present your ideas to others in a way that encourages rather than cuts off discussion. I decided to make clear these notes were first words, not last words and were intended to encourage conversation.

With that in mind, I jotted down the clerical aspects: each note should have a title, date, author institution, and number. I said the numbers would be handed out quickly upon request after the note was written so as to avoid holes in the sequence. And to emphasize the informal status of these notes, I said each of them, no matter what its content, would be called a Request for Comments or RFC². I stated that the content of an RFC could be "any thought, suggestion, etc., related to the HOST software or other aspect of the network." I added that an RFC should have at least one sentence and that it was more important for the notes to "be timely rather than polished." Then I added that the notes could contain "philosophical positions without examples," or "implementation techniques without introductory or background explication," or "questions without any attempted answers."³

Finally, I wrote a few sentences in an effort to explain what we were trying to do. I was hoping to avoid outside criticism and encourage a wide discussion, I opened with the phrase. "These standards (or lack of them)," were stated "explicitly for two reasons." First, I wanted to avoid the idea that these notes were standards or authoritative design documents but were intended "to promote the exchange and discussion of considerably less than authoritative ideas." Second, I felt it was important to bring forward incomplete or unpolished ideas and I wanted to ease the natural inhibition against doing so⁴.

I had expected that the RFCs would be temporary, probably replaced by more formal documentation when the network was up and running. To my surprise, the idea took hold, and RFCs persist to this day, albeit with major changes; the IETF's protocol standards are still published as RFCs. Because RFCs are online and available to anyone without cost, they form a powerful technical repository that has enabled generations of developers to extend the capabilities of the network in every imaginable direction.

The RFCs captured and represented many of the lessons that we learned in those first months developing the Arpanet. They were distributed via regular paper ("snail") mail, of course; successive versions of the recipient list for the RFCs were also distributed as RFCs. In a simple and practical sense, we had formed a network of people even before we had a working computer network⁵.

Our technical work had two tracks. We were slowly reaching toward a useful and workable set of abstractions as the basis for our protocols. We had to figure out how to connect each of our machines to its IMP, which did not feel like lofty research but was nonetheless challenging. We had to build both a hardware interface and an addition to the operating system. Since most of our machines were commercial products, changes to the hardware and modifications of the operating system were not common. The vendors who sold us the machines certainly did not expect such changes. At UCLA, the vendor's quote for a hardware interface was both too expensive and required too much time. Fortunately, Mike Wingfield, another of our graduate students, volunteered to do it less expensively and quite quickly. Others of us made the incisions into the operating system to make the Arpanet appear as a device available to user processes.

In a first test of the network, before we had completed the definition of the early protocols, we lashed up a connection between UCLA and SRI on October 29, 1969. The UCLA end acted as a user on a terminal connecting to SRI. And of course, when we tried to log into the SRI computer that day, we uncovered a bug which caused the SRI system to crash. But our work progressed steadily. By the start of 1970, new IMPs were added monthly and the communication between the IMPs was reasonably stable. We had learned how to network in all senses of the word—how to build network software and hardware, how to collaborate at a distance, and how to work in groups without well-defined authority. This experience became the dominant mode of protocol design, first for the Arpanet and then for the Internet. These methods were captured by the long-standing IETF rallying cry of "Rough consensus and running code" attributable to Dave Clark, derived from the dedicated efforts of that early group of graduate students who were learning how to network.

REFERENCES

- Page 6 of Oral History Interview with Frank Heart, Conducted by Judy O'Neill, March 13, 1990, Charles Babbage Institute, University of Minnesota, Minneapolis, conservancy.umn.edu/handle/11299/107349
- 2. The term RFC is now in the Oxford English Dictionary, credited to me. Recently, I learned Bill Duvall said he suggested the term. I do not recall, but all the more credit to him if so because the phrase seemed to me to strike just the right tone. Forty years later, I wrote an Op-Ed in the New York Times on the fortieth anniversary of the RFCs, "How the Internet Got Its Rules," https://www.nytimes.com/2009/04/07/opinion/ 07crocker.html

- Crocker, Steve, RFC 003, April 1969, tools.ietf.org/ html/rfc3
- 4. Ibid.
- 5. It is perhaps a reflection of our thinking that the network would facilitate distributed instead of centralized coordination that we instinctively required each RFC author to send copies to everyone listed on the current recipient list instead of having one particular site receive and then redistribute the RFCs.

Steve Crocker is an Internet pioneer and computer scientist with experience in academia, government, and industry. He managed research at DARPA in the early 1970s, founded the computer science research laboratory at the Aerospace Corporation in the early 1980s, co-founded CyberCash, Inc. in the early 1990s, and served on the board of the Internet Corporation for Assigned Names and Numbers from 2003 to 2017. He earned his B.A. in mathematics and Ph.D. in computer science from UCLA. His honors include the 2002 IEEE Internet Award, an honorary doctorate from the University of San Martin des Porres, Lima, Perú, founding membership in the Internet Hall of Fame Pioneers, and a Lifetime Achievement Award from the Internet Society. Contact the author at steve@shinkuro.com.

This article originally appeared in IEEE Annals of the History of Computing, *vol. 41, no. 2, 2019.*



IEEE Computer Graphics and Applications bridges the theory and practice of computer graphics. Subscribe to CG&A and

- stay current on the latest tools and applications and gain invaluable practical and research knowledge,
- discover cutting-edge applications and learn more about the latest techniques, and
- benefit from CG&A's active and connected editorial board.

Department

Supervised Learning for Fake News Detection

Julio C. S. Reis, André Correia, Fabrício Murai, Adriano Veloso, and Fabrício Benevenuto Universidade Federal de Minas Gerais

Editor: Erik Cambria, Nanyang Technological University, Singapore

Abstract—A large body of recent works has focused on understanding and detecting fake news stories that are disseminated on social media. To accomplish this goal, these works explore several types of features extracted from news stories, including source and posts from social media. In addition to exploring the main features proposed in the literature for fake news detection, we present a new set of features and measure the prediction performance of current approaches and features for automatic detection of fake news. Our results reveal interesting findings on the usefulness and importance of features for detecting false news. Finally, we discuss how fake news detection approaches can be used in the practice, highlighting challenges and opportunities.

Social media systems have been dramatically changing the way news is produced, disseminated, and consumed, opening unforeseen opportunities, but also creating complex challenges. A key problem today is that social media has become a place for campaigns of misinformation that affect the credibility of the entire news ecosystem.

A unique characteristic of news on social media is that anyone can register as a news publisher without any upfront cost (e.g., anyone can create a Facebook page claiming to be a newspaper or news media organization). Consequently,

Digital Object Identifier 10.1109/MIS.2019.2899143 Date of current version 3 May 2019. not only traditional news, corporations are increasingly migrating to social media (https:// www.comscore.com/Insights/Blog/Traditional-News-Publishers-Take-Non-Traditional-Path-to-Digital-Growth). Along with this transition, not surprisingly, there are growing concerns about fake news publishers posting "fake" news stories, and often disseminating them widely using "fake" followers.¹ As the extensive spread of fake news can have a serious negative impact on individuals and society, the lack of scalable fact checking strategies is especially worrisome.

Not surprisingly, recent research efforts are devoted not only to better comprehend this phenomenon¹ but also to automatize the detection of fake news.^{2,3,4} While a fully automated approach for the fake news problem can be quite

controversial and is still open for debate, a pertinent research question is: *What is the prediction performance of current approaches and features for automatic detection of fake news?*

Most of the existing efforts in this space are concurrent work, which identify recurrent patterns on fake news after they are already disseminated, or propose new features for training classifiers, based on ideas that have not been tested in combination. Thus, it is difficult to gauge the potential that supervised models trained from features proposed in recent studies have for detecting fake news. This paper briefly surveys existing studies on this topic, identifying the main features proposed for this task. We implement these features and test the effectiveness of a variety of supervised learning classifiers when distinguishing fake from real stories on a large, recently released and fully labeled dataset. Finally, we discuss how supervised learning models can be used to assist fact-checkers in evaluating digital content and reaching warranted conclusions.

FEATURES FOR FAKE NEWS DETECTION

Most of the existing efforts to detect fake news propose features that leverage information present in a specific dataset. In contrast, we use a recently released dataset that allows us to implement most of the proposed features explored in previous works.⁵ It consists of 2282 BuzzFeed news articles related to the 2016 U.S. election labeled by journalists and enriched with comments associated with the news stories as well as shares and reactions from Facebook users.

In this paper, we discarded stories labeled as "non factual content" (12%), and merged those labeled as "mostly false" (4%) and "mixture of true and false" (11%) into a single class, henceforth referred as "fake news." The remaining stories correspond to the "true" portion (73%). The rationale is that stories that mix true and false facts may represent attempts to mislead readers. Thus, we focus our analysis on understanding how features can be used to discriminate true and fake news.

On a coarse-grained level, features for fake news detection can be roughly categorized as follows: 1) features extracted from news content (e.g., language processing techniques); 2) features extracted from news source (e.g., reliability and trustworthiness); and 3) features extracted from environment (e.g., social network structure). Next, we briefly survey previous efforts, describing existing features and how we implemented them.

Textual Features consist of the information extracted from the news text, including the text body, the headline, and the text message used by the news source. For news articles embedded in images and videos, we applied image processing techniques for extracting the text shown on them. In total, we evaluated 141 textual features. Features were grouped in sets, which are described next.

- Language Features (Syntax): Sentence-level features, including bag-of-words approaches, "n-grams" and part-of-speech (POS tagging) were explored in previous efforts as features for fake news detection.^{2,6} Here, we implemented 31 features from this set including number of words and syllables per sentence as well as tags of word categories (such as noun, verb, adjective). In addition, to evaluate writers' style as potential indicators of text quality, we also implemented features based on text readability.
- 2) *Lexical Features:* Typical lexical features include character and word-level signals,^{7,6} such as amount of unique words and their frequency in the text. We implemented linguistic features, including number of words, first-person pronouns, demonstrative pronouns, verbs, hashtags, all punctuations counts, etc.
- 3) *Psycholinguistic Features:* Linguistic Inquiry and Word Count (LIWC)⁸ is a dictionarybased text mining software whose output has been explored in many classification tasks, including fake news detection.⁴ We use its latest version (2015) to extract 44 features that capture additional signals of persuasive and biased language.
- 4) Semantic Features: There are features that capture the semantic aspects of a text^{2,3} are useful to infer patterns of meaning from data.⁹ As part of this set of features, we consider the toxicity score obtained from Google's API (https://www.perspectiveapi.com/#/). The API uses machine learning models to quantify the extent to which a text (or comment, for instance) can be perceived as "toxic." We did

not consider strategies for topic extraction since the dataset used in this paper was built based on news articles about the same topic or category (i.e., politics).

5) *Subjectivity:* Using TextBlob's API (http://textblob.readthedocs.io/en/dev/), we compute subjectivity and sentiment scores of a text as explored in previous efforts.⁴

News Source Features consist of information about the publisher of the news article. To extract these features, we first parsed all news URLs and extracted the domain information. When the URL was unavailable, we associated the official URL of news outlet with news article. Therefore, we extract eight (eight) indicators of political bias, credibility and source trustworthiness, and use them as detailed next. Moreover, in this category, we introduce a new set composed of five features, called domain localization (see below).

- Bias: The correlation between political polarization and spread of misinformation was explored in previous studies.¹⁰ In this paper, we use the political biases of news outlets from the BuzzFeed dataset as a feature.
- 2) Credibility and Trustworthiness: In this feature set, we introduce seven new features to capture aspects of credibility (or popularity) and trustworthiness of domains. We collect, using Facebook's API (https://developers.facebook. com), user engagement metrics of Facebook pages that published news articles (i.e., "page talking about" count and "page fan" count). Then, we use the Alexa's API to get the relative position of news domain on the Alexa Ranking (https://www.alexa.com). Furthermore, using this same API, we collect Alexa's top 500 newspapers. Based on the intuition that some unreliable domains may try to disguise themselves using domains similar to those of well-known newspapers, we define the dissimilarity between domains from the Alexa ranking and news domains in our dataset (measured by the minimum edit distance) as features. Finally, we use indicators of low credibility of domains compiled¹¹ as features.
- 3) *Domain Location:* Ever since creating fake news became a profitable job, some cities

have become famous because of residents who create and disseminate fake news (https://www.bbc.com/news/magazine-

38168281). In order to exploit the information that domain location could carry, a pipeline was built to take each news website URL and extract new features, such as IP, latitude, longitude, city, and country. First, for each domain, the corresponding IP was extracted using the trace route tool. Then, the ipstack API was used to retrieve the location features. Although localization information (i.e., IP) has been previously used in works on bots or spam detection, to the best of our knowledge, there are no works that leverage these data in the context of fake news detection.

Environment Features consist of statistics of user engagement and temporal patterns from social media (i.e., Facebook). These features have been extensively used in previous efforts, ¹² especially to better understand the phenomenon of fake news. ¹³ Next, we detail the 21 features from this category.

- Engagement: We consider number of likes, shares, and comments from Facebook users. Moreover, we compute the number of comments within intervals from publication time (900, 1800, 2700, 3600, 7200, 14400, 28 800, 57 600 and 86 400 s), summing up to 12 features.
- 2) *Temporal Patterns:* Finally, to capture temporal patterns from user commenting activities, we compute the rate at which comments are posted for the same time windows defined before.

CLASSIFICATION RESULTS

We evaluate the discriminative power of the previous features using several classic and stateof-the-art classifiers, including *k*-Nearest Neighbors (KNN), Naive Bayes (NB), Random Forests (RF), Support Vector Machine with RBF kernel (SVM), and XGBoost (XGB). Given that we used hand-crafted features, there was no need to include a neural network model in the comparison since it would only associate weights with the features, rather than find new ones.

Classifier	AUC	F1
KNN	$0.80{\pm}0.009$	$0.75{\pm}0.008$
NB	$0.72{\pm}0.009$	$0.75 {\pm} 0.001$
RF	0.85±0.007	0.81±0.008
SVM	$0.79{\pm}0.030$	$0.76{\pm}0.019$
XGB	0.86±0.006	0.81±0.011

 Table 1. Results obtained for different classifiers w.r.t

 AUC and F1 score.

RF and XGB performed best.

We measure the effectiveness of each classifier w.r.t. the area under the ROC curve (AUC) and the Macro F1 score. In this case, the resulting AUC is the probability that a model will rank a randomly chosen fake news higher (more false) than a randomly chosen news article. The AUC is especially relevant for fake news detection since the decision threshold can be used to control the tradeoff between true and false positive rates. The F1 score combines precision and recall per class in a single metric and the Macro F1 score provides the overall performance of the classifier.

We compute 95% confidence intervals for the mean AUC and F1 by performing a fivefold split between training and test set, repeated ten times with different shuffled versions of the original dataset (a total of 50 runs). Table 1 shows the empirical results obtained from the fitted models using all features previously described.



Figure 1. ROC curve for the XGboost classifier. For BuzzFace, it is possible to correctly classify almost all of fake news with only 40% of false positive rate.



Figure 2. Chi-Square feature importance.

For each classifier, we learn a model from a set of previously labeled (i.e., preclassified) data, and then use it to classify new (unseen) news articles into "fake" or "not fake." The best results were obtained by RF and XGB classifiers, statistically tied with 0.85 (\pm 0.007) and 0.86 (\pm 0.006) for AUC, respectively.

Moreover, inspecting the ROC curve for XGB (see Fig. 1), we observe that it is possible to choose a threshold so as to correctly classify almost all of fake news (true positive rate \approx 1), while misclassifying 40% of the true news (false positive rate \approx 0.4). This can be useful, especially in assisting fact checkers to identify stories that are worth investigating. Finally, we assessed the relative power of the selected attributes in discriminating each class from the other by ranking features from each set based on X^2 (Chi Squared). Fig. 2 shows the results. Although all feature sets have some discriminatory power, there are some of them (e.g., credibility and localization of news sources, and news engagement) that can be more useful to improve the performance of models for fake news detection.

FAKE NEWS DETECTION IN PRACTICE

Fact checking is a damage control strategy that is both essential and not scalable. It might be hard to take out the human component out of the picture any time soon, especially if these news regard sensitive subjects such as politics. In the case of social networks and search engines, predictions made by models for fake news detection could be used internally to limit the audience of news stories likely to be fake. This is why automatic labeling of news stories raises so many questions about fairness and algorithm transparency, suggesting that it is likely that the final call will still depend on an expert at the end point for a long time.

On the bright side, automatic fake news detection could be used by fact checkers as an auxiliary tool for identifying content that is more likely to be fake. Our results show that the prediction performance of proposed features combined with existing classifiers has a useful degree of discriminative power for detecting fake news. Our best classification results can correctly detect nearly all fake news in our data, while misclassifying about 40% of true news, which is already sufficient to help fact checkers. In this context, providing explanations that supported the algorithm's output is crucial. For example, a certain story was considered false because it was posted by new newspaper hosted in the same IP address than a known blacklisted fake news source. Additionally, this kind of approach requires a continual pipeline where more stories get labeled each day and are, in turn, fed back to the models. Rather than verifying only the most suspicious stories, an active learning solution can be put in place, so that the model can also indicate which stories should be investigated in order to improve its prediction performance. More importantly, fake news is a relatively recent problem and the cost to label large datasets is still very high. In the future, larger volumes of labeled data will enable us to explore other techniques such as deep learning and push the boundaries of prediction performance.

ACKNOWLEDGMENTS

This work was supported in part by Google, CAPES, MASWeb (Grant FAPEMIG/PRONEX APQ-01400-14), CNPq, and Fapemig.

REFERENCES

- D. M. J. Lazer *et al.*, "The science of fake news," *Science*, vol. 359, no. 6380, pp. 1094–1096, 2018.
- N. J. Conroy, V. L. Rubin, and Y. Chen, "Automatic deception detection: Methods for finding fake news,"

in Proc. Annu. Meeting Assoc. Inf. Sci. Technol., 2015, pp. 1–4.

- W. Y. Wang, "Liar, liar pants on fire: A new benchmark dataset for fake news detection," in *Proc. Annu. Meeting Assoc. Comput. Linguistics*, 2017, pp. 422–426.
- S. Volkova, K. Shaffer, J. Jang Yea, and N. Hodas, "Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts on twitter," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, 2017, pp. 647–653.
- G. Santia and J. Williams, "BuzzFace: A news veracity dataset with facebook user commentary and egos," in *Proc. 12th Int. AAAI Conf. Web Soc. Media*, 2018, pp. 531–540.
- K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations Newslett.*, vol. 19, no. 1, pp. 22–36, 2017.
- C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *Proc. 20th Int. Conf. World Wide Web*, 2011, pp. 675–684.
- J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic inquiry and word count: LIWC 2001," Mahway: Lawrence Erlbaum Associates, vol. 71, 2001.
- E. Cambria, S. Poria, A. Gelbukh, and M. Thelwall, "Sentiment analysis is a big suitcase," *IEEE Intell. Syst.*, vol. 32, no. 6, pp. 74–80, Nov./Dec. 2017.
- F. N. Ribeiro, L. Henrique, F. Benevenuto, A. Chakraborty, J. Kulshrestha, M. Babaei, and K. P. Gummadi, "Media bias monitor: Quantifying biases of social media news outlets at large-scale.," in *Proc. of the Twelfth International AAAI Conference on Web and Social Media*, 2018, pp 290–299.
- C. Shao, G. L. Ciampaglia, O. Varol, A. Flammini, and F. Menczer, "The spread of low-credibility content by social bots," 2017, arXiv:1707.07592.
- M. Ebrahimi, A. H. Yazdavar, and A. Sheth, "Challenges of sentiment analysis for dynamic events," *IEEE Intell. Syst.*, vol. 32, no. 5, pp. 70–75, Sep./Oct. 2017.
- S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

Julio C. S. Reis is currently working toward the PhD degree in computer science at the Universidade Federal de Minas Gerais, Brazil. Contact him at julio. reis@dcc.ufmg.br.

André Correia is currently working toward the B.Sc. degree in information systems at the Universidade Federal de Minas Gerais, Brazil. His main interest is applied machine learning. Contact him at andrecorreia.dcc@gmail.com.

Fabrício Murai is an assistant professor in the Computer Science Department, Universidade Federal de Minas Gerais, Brazil. His research lies in the application of mathematical modeling, statistics and machine learning to informational and social networks. Contact him at murai@dcc.ufmg.br. **Adriano Veloso** is an associate professor of Computer Science at the Universidade Federal de Minas Gerais, Brazil. His interests are in machine learning and natural language processing. Contact him at adrianov@dcc.ufmg.br.

Fabrício Benevenuto is an associate professor in the Computer Science Department, Universidade Federal de Minas Gerais, Brazil. His research lies in topics related to social computing, computational journalism, and sentiment analysis. He is the corresponding author. Contact him at fabricio@dcc.ufmg.br.

This article originally appeared in IEEE Intelligent Systems, *vol. 34, no. 2, 2019.*

ADVERTISER INFORMATION

Advertising Personnel

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

Advertising Sales Representatives (display)

Central, Northwest, Southeast, 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

Advertising Sales Representative (Classifieds & Jobs Board)

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

Advertising Sales Representative (Jobs Board)

Marie Thompson Email: marie@4caradio.org Phone: 714-813-5094

CYBERTRUST



Thoughts on Cyberbullying

Nir Kshetri, University of North Carolina at Greensboro Jeffrey Voas, IEEE Fellow

Cyberbullying is a cybertrust issue that does not get much attention until after an incident occurs. It is hard to tie it directly to security or privacy. It is its own threat category, fitting between security, safety, and privacy. It is a social and societal issue, a cowardly action attempting to hide behind a virtual shield. Fortunately, digital forensics can usually unmask this.



seven times more likely than other people to be offenders.² In a survey conducted in 28 countries, 51% reported that the offenders in cyberbullying are classmates of cyberbullied children. The proportion was the highest in North America (65%) and the lowest in the Middle East/ Africa (39%).³ In most developed countries, cyberbullying is regarded as an important issue. In the United States, First Lady Melania Trump has made it a focus of her initiatives.

yberbullying is a targeted online weapon used by online offenders to inflict psychological and emotional harm to Internet users. According to a Pew Research Center survey from September 2018, 59% of U.S. teens had been bullied or harassed online.¹ Often, the offenders are family, friends, and other persons that the victims know and trust. For example, in cyberbullying incidents where children and young adults are victims, friends or dating partners are

Digital Object Identifier 10.1109/MC.2019.2898720 Date of publication: 16 April 2019 The seriousness of this issue has led to the emergence of new forms of cyberinsurance to protect against cyberbullies. Some insurers, such as American International Group (AIG) and the Arbella Insurance Group, have coverage options for cyberbullying, including coverage for costs incurred after a cyberbullying attack, e.g., legal expenses, temporary relocation expenses, and private tutoring. The Family CyberEdge policy, a new product from AIG, includes coverage for one year of psychiatric services if a family member is victimized by cyberbullying. Lost salary is also covered if the victim loses a job within 60 days.

46

GEOGRAPHIC, DEMO-GRAPHIC, ECONOMIC, AND SOCIOCULTURAL FACTORS

There are geographic, demographic, economic, and sociocultural variations in the awareness, patterns, and prevalence of cyberbullying. For instance, according to a study conducted in 28 countries by the global market research and consulting firm Ipsos, 75% of adults were aware of cyberbullying. However, the awareness varied from the highest in Sweden and Italy (91% each) and lowest in Saudi Arabia (37%).³

In the Ipsos study, 65% of parents reported that they know that this behavior takes place on social-networking sites.³ This proportion was higher in Latin America and the lowest in the Asia-Pacific region (Table 1).

There are also gender and economic dimensions of cyberbullying. For instance, girls are more likely to be victims of cyberbullying than boys, and poor children are more likely to be victims of cyberbullying than children from wealthier families.¹ There are also racial differences in terms of the perception of the seriousness of cyberbullying. For instance, in the United States, parents in general regard cyberbullying as among their top concerns related to their children's health and well-being. Among African-American parents specifically, however, cyberbullying was a relatively lower concern when compared to other social issues (Table 1).

ADDRESSING CYBERBULLYING

So what can be done to protect children from the psychological and emotional harm that results from cyberbullying? Table 2 shows coordinated efforts at various levels.

Fortunately, laws dealing with cyberbullying are evolving. The European Union's (EU's) General Data Protection Regulation (GDPR) has provisions that aim to protect children from cyberbullying and other misuse of information by social-media websites. The preamble to the GDPR states that children are "less aware of risks, consequences, safeguards, and their rights" related to personal data and often aren't able to take measures to control what is likely to happen with their personal data. The GDPR emphasizes the roles of parents and parental consent.⁴ Social networking sites require parental consent before they process children's information.⁵ In practice, this could mean that those under 16 years old may need to obtain their parents' permission to use social media.⁶ Individual EU member states can also lower the age required for parental consent from 16 years to as low as 13.⁷

In the United States, the 50 states, the District of Columbia, and U.S. territories have each taken various regulatory measures to addresses bullying in general and cyberbullying in particular.⁸ Some parents of cyberbullying victims have filed lawsuits against alleged bullies or schools for failing to protect their children. In early 2018, a Pennsylvania family sued Sean Davis, a player on the Pittsburgh Steelers football team, for cyberbullying the family's teenage son. The family accused Davis of posting a video on the social-media platform Snapchat that mocked the teen's work at a Chick-fil-A drive-through.⁹

TABLE 1. Geographic, demographic, economic, and sociocultural factors linked to cyberbullying.

Factor	Findings
Geography	Ipsos' Global Advisor study was conducted in 28 countries. The proportion of parents who reported that their own children or other children they knew were cyberbullied and that the harassing behavior took place on social- networking sites was 76% in Latin America compared to 53% in the Asia-Pacific region. ³
Gender	A Pew Research Center survey ¹ shows that, in the United States, 39% of girls were reported to be victims of false rumors online compared to 26% of boys; 29% of girls reported that they received unwanted explicit images compared with 20% of boys, and 15% of teen girls had become targets of four or more different forms of cyberbullying compared with 6% of boys. ¹ In India, about 90% of cyberstalking victims are women. ¹² In the Democratic Republic of Congo, women and LGBT groups are frequently targeted by cyberbullies. ¹⁵
Economic	According to Ipsos' Global Advisor study, in the United States, 24% of teens from families with annual household income lower than US\$30,000 a year had been the target of physical threats online, as compared with 12% of those with annual household income of US\$75,000 or more. ³
Race	According to a national U.S. survey of C.S. Mott Children's Hospital on Children's Health, University of Michigan, bullying and cyberbullying were the most serious concerns parents had about their children's health, followed by Internet safety. For African-American parents, racial inequities and school violence were bigger concerns. ¹⁶

CYBERTRUST

Many developing countries, on the other hand, lack laws that criminalize cyberbullying. As of 2016, China and Russia had no specific laws against cyberbullying. In Russia, cyberbullying is theoretically covered by conventional laws against violence or murder.¹⁰

Many other developing countries lack even basic cybercrime laws. According to a November 2016 report of the African Union Commission and the cybersecurity firm Symantec, only 11 countries in Africa had specific laws and provisions in place to deal with cybercrime and electronic evidence. An additional 12 countries had taken at least some legislative measures, albeit limited. In May 2018, Kenya enacted a law that criminalizes cyberbullying.¹¹ The Democratic Republic of Congo has no laws to protect people against cyberbullying. Prosecution for such offenses is nonexistent.¹⁰ Enactment of laws that criminalize cyberbullying is important to combat this problem.

There are also law-enforcement challenges in addressing these offenses. In India, the unsupportive attitudes of law-enforcement agencies and their unwillingness to help victims have contributed to a low reporting rate of cyberbullying cases.¹² To fight this problem, law enforcement must be better prepared.

Possible actions Actors Examples In Michigan, a "pattern of repeated harassment" is a felony Regulators Pass stricter laws that might discourage cyberbullying activities. that carries a penalty of up to five years in prison and a US\$5,000 fine.¹⁷ In The Netherlands, cyber offenders who engage in cyberbullying and harassment may face a prison sentence of up to 10 years. In 2018, a Dutch appeals court upheld this maximum prison sentence for a convicted cyberbully.²⁰ Law-enforcement Train law enforcement. In Illinois, police officers assigned to protect schools are required to undergo training focused on cyberbullying.²¹ agencies Organizations Educate students about cyberbullying's Seattle Public Schools participated in a pilot program and educational psychological and legal implications and with iCanHelpLine.org where subscribers can discuss institutions present to them actual case studies of issues related to student cyberbullying on social media. cyberbullying.²² iCanHelpline.org works with social media organizations, such as Instagram, Snapchat, and Twitter, to delete Invest in technical solutions, such as content.18 monitoring or blocking software to detect cyberbullying activities on school networks.²³ Technology Develop advanced technical tools. Credit report and identity theft protection company Identity Guard uses artificial intelligence to monitor social media feeds companies and identify behavior that can be considered cyberbullying. It uses IBM Watson to enable natural language processing and natural language classifiers. Complex algorithms identify potential cyberbullying instances and send alerts to parents. These alerts also include screenshots with dates and times of related content that triggered the warnings. Parents are then guided to resources, such as relevant laws and school policies, so that they can respond effectively.²⁴ Parents, guardians, Discuss Internet and cell phone etiquette. A survey of parents of teenagers found that more than 75% of parents discussed cyberbullying with their children, 86% and caregivers Talk with children about cyberbullying. joined their children's online social network to monitor interactions, and 67% monitored the security settings on their children's social media accounts.²⁵ The nonprofit organization End to Cyberbullying (ETCB) has Social arbiters (e.g., Create awareness about cyberbullying. press, governance Encourage victims to report abuses. taken initiatives to raise awareness about cyberbullying. It watchdog groups, works with students, educators, and parents. The ETCB has academics. and hundreds of volunteers worldwide.¹⁹ activists) The Family Online Safety Institute (FOSI) works with the industry, government, and other nonprofit organizations to address problems related to cyberbullying. FOSI also uses forums, conferences, special events, and YouTube to promote online safety.18

TABLE 2. Measures to address cyberbullying.

Parents and caregivers can also play a key role in helping children deal with cyberbullying. A study found that "authoritative" parents who listen to their children and provide guidance can help reduce the impact of cyberbullying.¹³

Cyberbullies as well as victims may be stigmatized, and victims may also suffer secondary victimization. Note that primary victimization occurs when a person becomes a victim of the act itself. Some mechanisms involved in primary victimization include physical/psychological suffering or financial losses. Secondary victimization takes place due to actions in the victim's social environment. Key mechanisms involved in secondary victimization include stigmatization, social isolation, and intrusive and humiliating questioning. Secondary victimization can also occur when journalists use faulty and insensitive practices in gathering or reporting news or when the criminal justice system takes inappropriate actions. Schools should have multiple avenues for reporting cyberbullying so that victims are not stigmatized.¹⁴

yberbullying can be as destructive as traditional bullying. However, parents often have a low level of awareness of this form of bullying. How people view and respond to this issue is shaped by various geographic, demographic, and racial differences. Constructive and supportive actions of parents can reduce the harm that cyberbullying causes. Organized and systematic responses by governments and law-enforcement agencies and by social arbiters, such as the press, governance watchdog groups, academics, and activists, can also play a role in fighting cyberbullying.

REFERENCES

 M. Anderson, "A majority of teens have experienced some form of cyberbullying," Pew Research Center,

- J. White, "How to keep cyberbullies out of your life," Inc., Oct. 5, 2017. Accessed on: Nov. 15, 2018. [Online]. Available: https://www.inc.com/john-white /how-to-keep-cyberbullies-out-ofyour-life.html
- M. Newall, "Cyberbullying: A global advisor survey," Ipsos, 2018.
 [Online]. Available: https://www .ipsos.com/sites/default/files/ct /news/documents/2018-06/cyber bullying_june2018.pdf
- 4. Better Internet for Kids, "Data Protection Directive or General Data Protection Regulation: Which one is for you?" Sept. 28, 2017. [Online]. Available: https://www.betterinternetforkids .eu/web/portal/practice/awareness /detail?articleId=2258465
- European Commission, "Can personal data about children be collected?" Accessed on: Nov. 15, 2018. [Online]. Available: https://ec.europa .eu/info/law/law-topic/data-protec tion/reform/rights-citizens /how-my-personal-data-protected /can-personal-data-about-childrenbe-collected_en
- V. Verdoodt, "Children's access to social media and the GDPR – 'Please mom, can I go on Facebook?"" Ku Leuven Centre for IT & IP Law, Aug. 9, 2016. [Online]. Available: https://www.law .kuleuven.be/citip/blog /childrens-access-to-social-media-andthe-gdpr-please-mom-can-i-go-onfacebook/
- J. Průša, "Beware of encouraging pupils to use social networks and messengers," CZ.NIC, Sept. 27, 2018. [Online]. Available: https://en.blog .nic.cz/2018/09/27/bewareof-encouraging-pupils-to-use-socialnetworks-and-messengers/
- Stopbullying.gov, "Laws & Policies," U.S. Department of Health and Human Services, Jan. 7, 2018. [Online]. Available: https://www.stopbullying .gov/laws/index.html

- 9. Associated Press, "Family sues Steelers' Sean Davis, alleging cyberbullying," USA Today, Feb. 27, 2018.
 [Online]. Available: https://www .usatoday.com/story/sports /nfl/steelers/2018/02/27/sean-davispittsburgh-steelers-cyberbullyinglawsuit/377666002/
- N. Kshetri, "Cybercrime and cybersecurity in India: Causes, consequences and implications for the future," Crime, Law Social Change, vol. 66, no. 3, pp. 313–338, 2016.
- A. Schwarz, "Kenya signs bill criminalising fake news," Mail & Guardian, May 16, 2018. [Online]. Available: https:// mg.co.za/article/2018-05-16-kenyasigns-bill-criminalising-fake-news
- P. K. Roy, "Why online harassment goes unpunished in India," BBC News, July 17, 2015. [Online]. Available: https://www.bbc.com/news /world-asia-india-33532706
- A. Baird, "Best defenses against cyber bullies," Scientific American, Aug. 24, 2010. [Online]. Available: https:// www.scientificamerican.com/arti cle/best-defenses-cyber-bullies/
- P. K. Smith and F. Thompson, "The best way to stop bullying: Get the cool kids to stick up for the victims," Washington Post, Aug. 8, 2014.
 [Online]. Available: https://www .washingtonpost.com/postevery thing/wp/2014/08/08/the-best-way-to-stop-bullying-in-schools/? utm_term=.bb405af86c8c
- K. Lyons, T. Phillips, S. Walker, J. Henley, P. Farrel, and M. Carpentier, "Online abuse: How different countries deal with it," *The Guardian*, Apr. 12, 2016. [Online]. Available: https:// www.theguardian.com/technology /2016/apr/12/online-abuse-ho w-harrassment-revenge-pornography-different-countries-deal-with-it
- News Medical Life Sciences, "Bullying and cyberbullying top parents' list of worries, new report reveals," Aug. 21, 2017. [Online]. Available: https://www .news-medical.net/news/20170821 /Bullying-and-cyberbullying-topparents-list-of-worries-new-reportreveals.aspx

- 17. L. Devito, "Cyberbullying is now a crime in Michigan punishable by jail time," *Detroit MetroTimes*, Dec. 28, 2018. [Online]. Available: https://www.metrotimes.com/news-hits/archives/2018/12/28/cyberbullying-is-now-a-crime-in-michigan-punish able-by-jail-time
- Seattle Public Schools, "District partners to stop bullying on social media," Sept. 29, 2017. [Online]. Available: https://www.seattleschools.org /district/calendars/news/what_s_new /district_partners_to_stop_bullying_ on_social_media
- E. Kaough, "Combatting cyberbullying: Government, NGO and the private sector," Ministry of Public Security, Israel. Accessed on: Nov.
 15, 2018. [Online]. Available: https:// www.gov.il/BlobFolder/reports /cyberbullying_brief/en/cyberbullying%20brief%2001.13.pdf
- 20. Associated Press, "Dutch court upholds maximum sentence for

cyberbully," 660 News, Dec. 14, 2018. [Online]. Available: https://www.660citynews.com /2018/12/14/dutch-court-upholdsmaximum-sentence-for-cyberbully

- NBC 5, "Illinois school resource officers to undergo training," NBC Universal, Aug. 20, 2018. [Online]. Available: https://www.nbcchicago.com /investigations/Illinois-School-Resource-Officers-to-Undergo-Training-491310131.html
- 22. M. Clifford, "15 strategies educators can use to stop cyberbullying." informED, Oct. 26, 2012. [Online]. Available: https://www .opencolleges.edu.au/informed /features/15-strategies-educatorscan-use-to-stop-cyberbullying/
- 23. C. Page, "Striking back at the cyberbullies," BBC News, Apr. 16, 2006.
 [Online]. Available: http://news
 .bbc.co.uk/2/hi/uk/4912766.stm
- 24. T. Meier, "AI technology helps protect teens from cyberbullying," IBM, Feb. 27,

This article originally appeared in Computer, *vol. 52, no. 4, 2019.*

2018. [Online]. Available: https://www .ibm.com/blogs/client-voices /ai-technology-protect-teenscyberbullying/

25. HealthDay, "Cyberbullying a big worry for parents: Survey," July 15, 2011. [Online]. Available: https:// consumer.healthday.com /health-technology-information-18 /misc-computer-health-news-150 /cyberbullying-a-big-worry-for-parentssurvey-654818.html

NIR KSHETRI is a professor of management in the Bryan School of Business and Economics at the University of North Carolina at Greensboro. Contact him at nbkshetr@uncg.edu.

JEFFREY VOAS is an IEEE Fellow, is Computer's "Cybertrust" column editor, and was a cofounder of Cigital. Contact him at j.voas@ieee.org.



CAREER OPPORTUNITIES

The Department of Computer and Cyber Sciences at the US Air Force Academy

seeks to fill a faculty position at the Assistant Professor level. Exceptionally qualified candidates at upper ranks will also be considered. The department is particularly interested in candidates with a background in cybersecurity, but all candidates with a passion for teaching computer science are encouraged to apply.

The Academy is a national service institution, charged with producing leaders of character for the US Air Force. Faculty members are expected to exemplify the highest ideals of professionalism and integrity. The Academy is located in Colorado Springs, an area known for its natural beauty and quality of life. The United States Air Force Academy values the benefits of diversity among the faculty to include a variety of educational backgrounds, professional and life experiences.

For information on how to apply, go to usajobs.gov and search with the keyword 545526600. US citizenship is required. Candidates with specific questions can contact Dr Barry Fagin at barry.fagin@usafa.edu.



Call for Articles





IEEE Software seeks practical, readable articles that will appeal to experts and nonexperts alike. The magazine aims to deliver reliable information to software developers and managers to help them stay on top of rapid technology change. Submissions must be original and no more than 4,700 words, including 250 words for each table and figure.

Author guidelines: www.computer.org/software/author Further details: software@computer.org www.computer.org/software

The University of Alabama in Huntsville

The Department of Computer Science at The University of Alabama in Huntsville (UAH) invites applicants for a tenure-track faculty position at the **Assistant Professor** level beginning August 2020 to support the gaming and entertainment computing program.

A Ph.D. in computer science or a closely related area is required. The successful candidate will have a strong academic background and be able to secure and perform

funded research in areas typical for publication in well-regarded academic conference and journal venues. In addition, the candidate should embrace the opportunity to provide undergraduate education.

The department has a strong commitment to excellence in teaching, research, and service; the candidate should have good communication skills, strong teaching potential, and research accomplishments.

UAH is located in an expanding, high-technology area, in close proximity to Cummings Research Park, the second largest research park in the nation and the fourth largest in the world. Nearby are the NASA Marshall Space Flight Center, the Army's Redstone Arsenal, numerous Fortune 500 and high tech companies. UAH also has an array of research centers, including information technology and cybersecurity. In short, collaborative research opportunities are abundant, and many well-educated and highly technically skilled people are in the area. There is also access to excellent public schools and inexpensive housing.

UAH has an enrollment of approximately 9,900 students. The Computer Science department offers BS, MS, and PhD degrees in Computer Science and contributes to interdisciplinary degrees. Faculty research interests are varied and include cybersecurity, mobile computing, data science, software engineering, visualization, graphics and game computing, multimedia, AI, image processing, pattern recognition, and distributed systems. Recent NSF figures indicate the university ranks 30th in the nation in overall federal research funding in computer science.

Interested parties must submit a detailed resume with references to <u>info@cs.uah.edu</u> or Chair, Search Committee, Department of Computer Science, The University of Alabama in Huntsville, Huntsville, AL 35899. Qualified female and minority candidates are encouraged to apply. Initial review of applicants will begin as they are received and continue until a suitable candidate is found.

The University of Alabama in Huntsville is an affirmative action/equal opportunity employer/minorities/ females/ veterans/ disabled.

Please refer to log number: 20/21-549

COMPUTER ARCHITECTURE

LETTERS

IEEE Computer Architecture Letters is a forum for fast publication of new, high-quality ideas in the form of short, critically refereed technical papers. Submissions are accepted on a continuing basis and letters will be published shortly after acceptance in IEEE Xplore and in the Computer Society Digital Library.

Submissions are welcomed on any topic in computer architecture, especially:

- Microprocessor and multiprocessor systems
- Microarchitecture and ILP processors
- Workload characterization
- Performance evaluation and simulation techniques
- Interactions with compilers and operating systems
- Interconnection network architectures
- Memory and cache systems
- Power and thermal issues at the architectural level
- I/O architectures and techniques
- Independent validation of previously published results
- Analysis of unsuccessful techniques
- Domain-specific processor architecture (embedded, graphics, network)
- High-availability architectures
- Reconfigurable computer architectures

www.computer.org/cal



Join the IEEE Computer Society for subscription discounts today!

www.computer.org/product/journals/cal



SHARE AND MANAGE YOUR RESEARCH DATA

IEEE DataPort is an accessible online platform that enables researchers to easily share, access, and manage datasets in one trusted location. The platform accepts all types of datasets, up to 2TB, and dataset uploads are currently free of charge.



IEEEDataPort[™] upload datasets at ieee-dataport.org

IEEE

LETTERS OF THE COMPUTER SOCIETY



IEEE Letters of the Computer Society (LOCS) is a rigorously peer-reviewed forum for rapid publication of brief articles describing high-impact results in all areas of interest to the IEEE Computer Society.

Topics include, but are not limited to:

- software engineering and design
- information technology
- software for IoT, embedded, and cyberphysical systems
- $\cdot \;$ cybersecurity and secure computing
- autonomous systems
- machine intelligence
- parallel and distributed software and algorithms
- programming environments and languages
- \cdot $\,$ computer graphics and visualization $\,$
- services computing
- databases and data-intensive computing
- cloud computing and enterprise systems
- hardware and software test technology

OPEN ACCESS

LOCS offers open access options for authors. Learn more about IEEE open access publishing:

https://open.ieee.org

EDITOR IN CHIEF

Darrell Long – University of California, Santa Cruz

ASSOCIATE EDITORS

- Sasitharan Balasubramaniam Waterford Institute of Technology and Tampere University
- Dirk Duellmann CERN
- Dan Feng Huazhong University of Science and Technology
- Gary Grider Los Alamos National Laboratory
- Kanchi Gopinath Indian Institute of Science (IISc), Bangalore
- James Hughes University of California, Santa Cruz
- Ilia Iliasdis IBM Research Zurich
- Katia Obraczka University of California, Santa Cruz
- Mubashir Husain Rehmani Cork Institute
 of Technology
- Thomas Johannes Emil Schwarz
 Marquette University
- Marc Shapiro Sorbonne-Université– LIP6 & Inria
- Kwang Mong Sim Shenzhen University

Submit / Subscribe / Learn More www.computer.org/locs







Keep up with the latest IEEE Computer Society publications and activities wherever you are.

Follow us:

f in

 (\mathbf{O})

- @ComputerSociety
- facebook.com/IEEEComputerSociety
- IEEE Computer Society
 - youtube.com/ieeecomputersociety
 - instagram.com/ieee_computer_society





SCOPE

The *IEEE Transactions on Big Data* (*TBD*) publishes peer reviewed articles with big data as the main focus. The articles provide cross disciplinary innovative research ideas and applications results for big data including novel theory, algorithms and applications. Research areas for big data include, but are not restricted to, big data analytics, big data visualization, big data curation and management, big data semantics, big data infrastructure, big data standards, big data performance analyses, intelligence from big data, scientific discovery from big data security, privacy, and legal issues specific to big data. Applications of big data in the fields of endeavor where massive data is generated are of particular interest.

SUBSCRIBE AND SUBMIT

For more information on paper submission, featured articles, calls for papers, and subscription links visit:







SECURITY& PRIVACY

IEEE Security & Privacy is a bimonthly magazine communicating advances in security, privacy, and dependability in a way that is useful to a broad section of the professional community.

The magazine provides articles with both a practical and research bent by the top thinkers in the field of security and privacy, along with case studies, surveys, tutorials, columns, and in-depth interviews. Topics include:

- Internet, software, hardware, and systems security
- Legal and ethical issues and privacy concerns
- Privacy-enhancing technologies
- Data analytics for security and privacy
- Usable security
- Integrated security design methods
- Security of critical infrastructures
- Pedagogical and curricular issues in security education
- Security issues in wireless and mobile networks
- Real-world cryptography
- Emerging technologies, operational resilience, and edge computing
- Cybercrime and forensics, and much more

www.computer.org/security



Join the IEEE Computer Society for subscription discounts today!







I EEE Computer Society conferences are valuable forums for learning on broad and dynamically shifting topics from within the computing profession. With over 200 conferences featuring leading experts and thought leaders, we have an event that is right for you.

Find a region:AfricaAustraliaNorth AmericaAsiaAsiaEuropeSouth America

NOVEMBER

4 November

ICTAI (IEEE 31st Int'l Conf. on Tools with Artificial Intelligence)

7 November

SEC (IEEE/ACM Symposium on Edge Computing)

8 November

• ICDM (IEEE Int'l Conf. on Data Mining) 🔺

9 November

 FOCS (IEEE 60th Annual Symposium on Foundations of Computer Science)

11 November

• ASE (34th IEEE/ACM Int'l Conf. on Automated Software Eng.) •

17 November

- ICCD (IEEE 37th Int'l Conf. on Computer Design) ▲
- SC19 (SC19: Int'l Conf. for High Performance Computing, Networking, Storage and Analysis)

18 November

• BIBM (IEEE Int'l Conf. on Bioinformatics and Biomedicine) •

DECEMBER

3 December

• RTSS (IEEE Real-Time Systems Symposium) 🔺

4 December

• IREHI (IEEE Int'l Rural and Elderly Health Informatics Conf.)

9 December

- AIVR (IEEE Int'l Conf. on Artificial Intelligence and Virtual Reality)
- Big Data (IEEE Int'l Conf. on Big Data) 🕨
- CDKE (IEEE Int'l Conf. on Conversational Data & Knowledge Eng.)
- ISM (IEEE Int'l Symposium on Multimedia) >

10 December

• ISSPIT (IEEE Int'l Symposium on Signal Processing and Information Technology) ▲

2020

January

13 January

ICCPS (Int'l Conf. on Cyber-Physical Systems)

February

3 February

ICSC (IEEE 14th Int'l Conf. on Semantic Computing)

18 February

SANER (IEEE 27th Int'l Conf. on Software Analysis, Evolution and Reengineering)

19 February

• BigComp (IEEE Int'l Conf. on Big Data and Smart Computing) ▲

22 February

• CGO (IEEE/ACM Int'l Symposium on Code Generation and Optimization) •

March

2 March

 WACV (IEEE Winter Conf. on Applications of Computer Vision)

9 March

- DATE (Design, Automation & Test in Europe Conf. & Exhibition) ●
- IRC (4th IEEE Int'l Conf. on Robotic Computing) ▲

16 March

• ICSA (IEEE Int'l Conf. on Software Architecture) ★

22 March

 VR (IEEE Conf. on Virtual Reality and 3D User Interfaces)

23 March

- ICST (13th IEEE Conf. on Software Testing, Validation and Verification) ●
- PerCom (IEEE Int'l Conf. on Pervasive Computing and Communications)

April

5 April

• ISPASS (Int'l Symposium on Performance Analysis of Systems and Software) •

14 April

- PacificVis (IEEE Pacific Visualization Symposium) ▲
- 20 April
- ICDE (IEEE 36th Int'l Conf. on Data Eng.) >

May

3 May

 FCCM (IEEE 28th Annual Int'l Symposium on Field-Programmable Custom Computing Machines)

4 May

• HOST (IEEE Int'l Symposium on Hardware Oriented Security and Trust) >

18 May

- SP (IEEE Symposium on Security and Privacy)
- FG (IEEE Int'l Conf. on Automatic Face and Gesture Recognition) ★
- IPDPS (IEEE Int'l Parallel and Distributed Processing Symposium) •

23 May

 ICSE (IEEE/ACM 42nd Int'l Conf. on Software Eng.) ▲

30 May

• ISCA (ACM/IEEE 47th Annual Int'l Symposium on Computer Architecture) ●

June

14 June

• CVPR (IEEE Conf. on Computer Vision and Pattern Analysis) •

16 June

• EuroS&P (IEEE European Symposium on Security & Privacy) (Location TBD)

19 June

 JCDL (ACM/IEEE Joint Conf. on Digital Libraries) ▲

29 June

• DSN (50th Annual IEEE/IFIP Int'l Conf. on Dependable Systems and Networks) ●

30 June

 MDM (21st IEEE Int'l Conf. on Mobile Data Management)

 Image: Management
 <li

Learn more about IEEE Computer Society Conferences

www.computer.org/conferences

IEEE Computer Society Kicks Off Cybersecurity Awareness Month

The IEEE Computer Society kicks off a month-long campaign of exciting activities that highlight cybersecurity awareness, encourage accountability, and promote proactive behavior to ensure best security practices for individuals and their organizations.

- Special blog posts
- Daily cybersecurity tips
- Security and Privacy TC Membership
- Cybersecurity curricular guidelines
- Cipher Newsletter
- And more!

Visit computer.org for daily updates!

Don't miss our lineup of exclusive activities and resources!

PODCAST Evan Gilman and Doug Barth **EPISODE** Securing Authors of Zero-Trust Networks: Building **SE RADIO** Your API Secure Systems in Untrusted Networks 22 OCTOBER PODCAST Neil Madden **EPISODE** Author of API Security in Action **8 OCTOBER** and Security Director **WEBINAR** Steven Bav of ForgeRock Former NSA 23 OCTOBER now leading author

SE RADIO





Zero-Trust Networks